N465

D

NAVAL RESEARCH LOGISTICS QUARTERLY

JUNE 1974
VOL. 21, NO. 2

OFFICE OF NAVAL RESEARCH

NAVSO P-1278

NAVAL RESEARCH LOGISTICS QUARTERLY

The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.

The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, June, September, and December and can be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Subscription Price: $10.00 a year in the U.S. and Canada, $12.50 elsewhere. Cost of individual issues may be obtained from the Superintendent of Documents.

OPTIMAL DESIGN OF A MULTI-ITEM, MULTI-LOCATION, MULTI-REPAIR TYPE REPAIR AND SUPPLY SYSTEM*

Evan L. Porteus

Graduate School of Business
Stanford University

and

Zachary F. Lansdowne

Control Analysis Corp.
Palo Alto, California

ABSTRACT

The design of a system with many locations, each with many items which may fail while in use, is considered. When items fail, they require repair; the particular type of repair being governed by a probability distribution. As repairs may be lengthy, spares are kept on hand to replace failed items. System ineffectiveness is measured by expected weighted shortages over all items and locations, in steady state. This can be reduced by either having more spares or shorter expected repair times. Design consists of a provisioning of the number of spares for each item, by location; and specifying the expected repair times for each type of repair, by item and location. The optimal design minimizes expected shortages within a budget constraint, which covers both (i) procurement of spares and (ii) procurement of equipment and manning levels for the repair facilities. All costs are assumed to be separable so that a Lagrangian approach is fruitful, yielding an implementable algorithm with outputs useful for sensitivity analysis. A numerical example is presented.

1. INTRODUCTION

Consider a multi-item, multi-location, multi-repair type supply and repair system, which, for a given item and location, can be diagramed as in Figure 1.

When fully operative, the system at that location employs n units of the item, called *operational units.* In addition to these n, s spares are retained in a pool, to be used to replace operational units, when they fail.

When a unit fails, one of the (operative) spares in the pool is used to replace the inoperative unit. If the pool of spares is exhausted, then a *shortage* (backorder) occurs, and the number of operational units decreases below n. Meanwhile, the nature of the failure of the unit is determined. There are K distinct types of failure possible, each with a corresponding type of repair required. These may range from tightening a screw to irreparable failure. (In the latter case, "repair" would consist of new procurement. We discuss this and related points in more detail later.) In the diagram p_k denotes the probability that type k repair is required.

Once the type of repair required is determined, the unit begins to undergo such repair. Once repair is completed, the unit is returned to the pool of spares. If a shortage exists at that point, the unit is immediately installed in the system.

*This research was sponsored by the Office of Naval Research under contract N00014–72–C–0266 (NR–347–022).

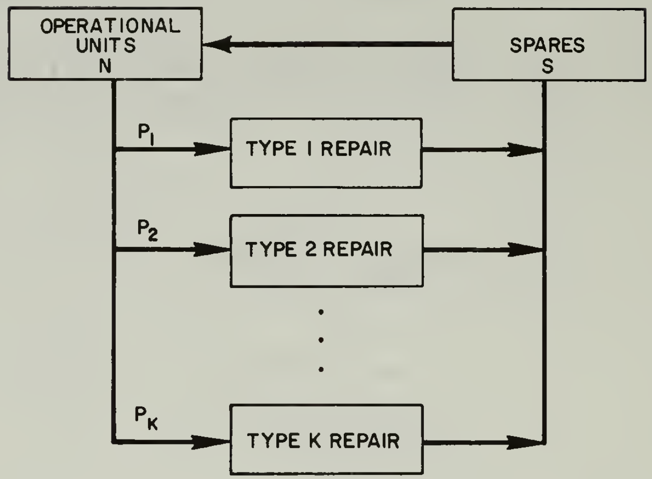

FIGURE 1. Diagram of a multi-repair-type repair and supply system for one item at one location.

System performance is measured by expected weighted shortages, in steady state, summed over all items and locations; called *expected shortages*, for short.

The design problem faced by the supply side of the system is to determine a provisioning of the number of spares for each item at each location. Adding spares reduces expected shortages while increasing cost. However, there are other ways of reducing expected shortages, namely by improving the performance of the repair side of the system.

Recall that once a unit has failed and the type of repair required determined, the unit remains inaccessible to the system during the *repair time*, consisting of the time elapsing from the point of failure to the point at which the unit is returned to the pool. Various components of a repair time are discussed later. This repair time is a random variable whose distribution function influences expected shortages. In our model, though, this influence will be determined entirely by the expected repair times for each type of repair. This point will be discussed in detail later. The design problem on the repair side of the system is then to determine the expected repair times for each type of repair, for each item, by location. Lowering expected repair times reduces expected shortages while increasing cost. The "overall" design problem is therefore to determine (i) how many spares of each item to provision to each location, and (ii) what the expected repair times ought to be, for each repair type, by item and location.

We seek an optimal design; that is, a design which minimizes expected shortages subject to a budget constraint covering total expenditures on spares and repair times. Possible relevant costs are identified in section 4.5, and considerations leading to expressing these expenditures on a common basis are also discussed.

With the exception of [10], previous studies have focused on either the supply side of the system (e.g., [14]) or the repair side (e.g., [9]), but not on both simultaneously. Our approach allows tradeoffs between these two sides to be made, in the interest of improving overall system performance. Our aim has been to provide the practitioner with a model which might be of immediate use in analyzing these tradeoffs. We have therefore made simplifying assumptions when they would simplify the model significantly. We highlight the important such assumptions at the end of the paper.

The next section describes the model for a given item at a given location. Section 3 then explicitly defines the optimization model, incorporating all items and locations. Section 4 follows with an interpretation for the use of the model. Model development occurs in the next five sections. Section 10 provides a numerical example, and section 11 ends the paper with a brief discussion of its relationship to previous work.

2. DESCRIPTION OF THE MODEL FOR A GIVEN ITEM AND LOCATION

We center attention on one item at one location. Let $N^F(t)$ denote the number of failures (of operational units) which have occurred during $[0, t]$ for $t \geq 0$. We call $\{N^F(t);\ t \geq 0\}$ the *failure process* and assume it is a Poisson process (e.g., [13]) with rate λ ($\lambda > 0$).

The design parameters under control are s, the number of spares, and the sequence $\{T_k;\ k=1, 2, \ldots, K\}$ of expected type k repair times.

For precision, assume that, for each repair type k, there is a single design parameter under control, denoted by T_k, and that $0 < T_k < \infty$.

Let $X_{kn}(T_k)$ denote the repair time for the nth unit to undergo type k repair, given T_k, for each k and n ($k=1, 2, \ldots, K$ and $n=1, 2, \ldots$). We assume

$$\{X_{11}(T_1), X_{12}(T_1), \ldots, X_{21}(T_2), X_{22}(T_2), \ldots, X_{K1}(T_K), X_{K2}(T_K), \ldots\},$$

are mutually independent random variables for all specification of $\{T_k\}$, and that, for each k,

$$\{X_{k1}(T_k), X_{k2}(T_k), \ldots\},$$

are identically distributed random variables, with distribution function F_k; i.e.,

$$F_k(x;\ T_k) \equiv P\{X_{kn}(T_k) \leq x\} \qquad \text{for each } k \text{ and } n.$$

We also assume that, for each k, the first moment of X_{kn} exists, is unique and equals T_k; i.e.,

$$\int_{x=0}^{\infty} x dF_k(x;\ T_k) = T_k \qquad \text{for } 0 < T_k < \infty.$$

Thus, the repair parameters under control are the expected repair times. We often suppress $\{T_k\}$ in the remainder of this section, as its specification remains fixed.

Let $Y_n = Y_n(\{T_k\})$ denote the repair time for the nth failure (the nth operational unit to fail and thereby require repair), for each n, (given the specification $\{T_k\}$). We assume $\{Y_1, Y_2, \ldots\}$ are independent, identically distributed random variables with distribution function F given by

$$F(x) \equiv \sum_{k=1}^{K} p_k F_k(x,\ T_k).$$

That is, given a failure, type k repair is required with probability p_k, for each k. Thus, (given T_k) the

expected repair time for the nth failure is independent of n and equals

$$T \equiv \sum_{k=1}^{K} p_k T_k.$$

Let $N^R(t)$ denote the total number of units in repair (of any type) at time $t \geqslant 0$. Let $N^S(t)$ denote the number of shortages at time $t \geqslant 0$. Clearly

$$N^S(t) = \begin{cases} 0 & \text{if } N^R(t) \leqslant s \\ N^R(t) - s & \text{otherwise.} \end{cases}$$

In short, $N^S(t) = [N^R(t) - s]^+$.

Let $S(t)$ $(=S(t; s, \{T_k\})) \equiv E[N^S(t)]$, the expected number of shortages at time $t \geqslant 0$. If it exists, let

$$S(s, \{T_k\}) \equiv \lim_{s \to \infty} S(t).$$

Let $c^S(s)$ denote the cost of retaining s spares in the system and $c_K^R(T_k)$ the cost of setting the expected type k repair time equal to T_k, for each k.

3. THE OPTIMIZATION PROBLEM OVER ALL ITEMS AND LOCATIONS

In this section, in contrast to the last, we indicate the notational dependence on the item i and location j under consideration. For example $c_{ij}^S(s_{ij})$ denotes the cost of retaining s_{ij} spares of item i at location j, for each i and j. The optimization problem can then be stated as: find s_{ij} and T_{ijk} for each i, j, and k to

$$\text{(1.1)} \qquad \min \sum_{ij} w_{ij} S_{ij}(s_{ij}, T_{ij1}, \ldots, T_{ijK})$$

s.t.

$$\text{(1.2)} \qquad \sum_{ij} c_{ij}^S(S_{ij}) + \sum_{ijk} c_{ijk}^R(T_{ijk}) \leqslant b,$$

where w_{ij} is a weight, assigned in relation to the importance of a shortage of item i at location j, and b is a real number denoting the budget, which applied to the total expenditures on spares and expected repair times. For notational convenience, we assume hereafter that $w_{ij} = 1$ for all i and j. The adjustments are easily made in the more general case.

Our approach to solving (1) (consisting of (1.1) and (1.2), collectively) is a Lagrangian one. That is, (1) is of the form: find x in R^n to

$$(P_\beta) \qquad \min f(x) \qquad \text{s.t.} \quad g(x) \leqslant \beta,$$

where $\beta \epsilon R$, and (1) corresponds to (P_b). We say x^* is *undominated* if x^* is optimal for (P_β) for some β. We form the Lagrangian $L(x, u) \equiv f(x) + u[g(x) - \beta]$ and employ the following:

LEMMA 1: (Everett [5], Brooks and Geoffrion [2]). If x^* minimizes $L(\cdot, u)$, (where $u \geq 0$), then x^* is undominated. In fact x^* is optimal for (P_β) if $u[g(x^*) - \beta] = 0$.

Thus, in particular, if x^* minimizes $L(\cdot, u)$ for some u and $g(x^*) = b$, then x^* is optimal for (1). The problem is to find an appropriate multiplier, if one exists.

There are several approaches one can take to finding such a multiplier. One is to use linear programming iteratively, as suggested by Brooks and Geoffrion [2]. As seen in Zangwill [20], this can be thought of as the Dantzig-Wolfe ([4] and [3, chap. 22]) method. For the single constraint problem we face, certain specializations and improvements can be made, as in Greenberg [8]. However, these approaches require the data for a given item and location to be accessed several times during the algorithm. In practice, there will likely be many item/location combinations, so that most of the data would have to be stored in slow access, peripheral storage. We therefore take the approach of Fox and Landi [6], whereby a sequence $u_1 < u_2 < \ldots < u_N$ of positive multipliers in specified. For each i, $L(x, u_i)$ is minimized, yielding $x(u_i)$, which is optimal for (P_β), where $\beta = (g(x(u_i)))$. It follows that $f(x(u_1)) \leq f(x(u_2)) \leq \ldots \leq f(x(u_N))$, and $g(x(u_1)) \geq g(x(u_2)) \geq \ldots \geq g(x(u_N))$, so that the best solution which is feasible for (P_b) is $x(u_i)$ where i is the smallest index such that $g(x(u_i)) \leq b$. This approach requires only one access to the data for each item/location combination. Another advantage is that it provides an optimal solution for each of several problems; ones with different budget levels. That is, each $x(u_i)$ is undominated. These results can often be an important aid in a sensitivity analysis of the relation between expected shortages and budget level.

A disadvantage of this method is that we must assume $g(x(u_N)) \leq b$, so that at least one of the undominated solutions is feasible for (P_b). Another disadvantage is that the grid of multipliers chosen may be too coarse, resulting in too wide a range of budget levels for which optimal solutions have been determined. In either case, another pass through the algorithm using a different (e.g., finer) grid might be required, and, indeed, the same difficulty might again manifest itself. A further difficulty is that there may be no multiplier which yields an optimal solution for a particular budget in mind. However, this problem is shared by all approaches using the Lagrangian as we have defined it. (See [7] for a more general Lagrangian which eliminates the current problem but adds others.) Nevertheless, in practice, once the model is put into implementation, a fairly good range of multipliers would likely be known. In addition, often no particular budget level is in mind. Indeed, this model might well be used to determine what the budget level *ought* to be, by showing the relationship between expected shortages and budget level. (However, if a particular budget level was in mind, and repeated access to the data for each item/location combination posed no computational burden, then Greenberg's [8] approach would be preferred.)

Note that the Lagrangian for (1) is separable by item and location. That is, let

$$L_{ij}(s_{ij}, T_{ij1}, \ldots, T_{ijK}, u) \equiv S_{ij}(s_{ij}, T_{ij1}, \ldots, T_{ijK}) + uc^S_{ij}(s_{ij}) + u \sum_k c^R_{ijk}(T_{ijk}).$$

Then we have

$$L(x, u) = \sum_{ij} L_{ij}(s_{ij}, T_{ij1}, \ldots, T_{ijk}, u) - ub.$$

Thus, given u, if we minimize each term separately, we minimize the sum. This justifies concentrating on a single item at a single location in the remaining section. That is, we shall consider the problem: for fixed u, find $s, T_1, \ldots, T_K$ to

$$\min L(s, T_1, \ldots, T_K, u), \tag{2}$$

where

$$L(s, T_1, \ldots, T_K, u) \equiv S(s, T_1, \ldots, T_K) + uc^S(s) + u \sum_k c_k^R(T_k).$$

4. INTERPRETATION OF THE MODEL

4.1. Repair Times as Response Times

Repair time usually consists, in practice, of the sum of the times elapsing while detecting and diagnosing a failure, transporting the unit to the appropriate repair facility, awaiting physical repair, further diagnosing for exact repairs needed, conducting physical repair (including, but not limited to replacing inoperative components), and transporting back to the pool of spares. Thus, changes in any of these times will affect "repair time," which some readers may therefore wish to call "response time."

4.2. Repair Times are I.I.D. Random Variables

In practice, a repair type might well be selected to represent a given repair facility, which would be characterized by its equipment, manning levels for different skill categories, inventories of components used in repairs, etc. Operation of this facility might be modeled by a multi-server queueing system, where a server might represent a given crew of men and equipment, with specialized capabilities. Thus, each of the servers might have different characteristics. In addition, since many different items would be sent to this facility to be repaired, service time (time to repair the item) would depend not only on the server, but on the item being served (repaired). Attempting to take all of this detail into account in our current model would likely make the model too cumbersome to be of practical use. To simplify things, we assume that, for each item and location, the repair times $\{X_{k1}(T_k), X_{k2}(T_k), \ldots\}$ are i.i.d. random variables. This would not be true in a detailed model of the type just discussed. It would be if the detailed model incorporated an infinite number of servers, and, if there were a large number of servers, such an assumption should yield a good approximation. However, in such a case, units sent to be repaired would begin physical repair as soon as they arrived at the repair facility. That is, they would never wait to begin service. This would rarely be the case in practice. However, this does not mean we assume items never wait to begin physical repair. Indeed, we would want to include an estimate of the distribution of the time awaiting physical repair in our estimate of the distribution of "repair time."

In short, our assumption, that all repair times are mutually independent and those for a particular item/location combination are also identically distributed, is a crucial one and there are several ways in which we might justify it, as an approximation. It allows us to decompose the problem into separate ones for each item/location combination, and this decomposition is critical in keeping the model tractable.

4.3. A Repair Type as a Supply Function

4.3.1. A Repair Type as a Joint Repair and Supply Function

We may select one of the repair types to represent a central repair facility which also maintains spares of each item. When a unit fails and requires central repair, then the central repair facility is notified that the unit is being sent for repair. Rather than waiting for the unit to arrive, to then repair it and send it back, the central facility immediately sends one of its spares, if it has one available. If none are available, the next unassigned one coming out of repair is sent.

In this case, expected "repair time" consists of the expected time to notify the central facility, time to locate an unassigned unit there, and time to transport it back to the pool of spares. The time to locate an unassigned unit at the central facility is negligible if operative spares are available there, but may consist of as much as the time required to transport the item from the initial location to the central facility, repair the unit there, and transport it back to the pool of spares (at the initial location).

This line of thought is due to Sherbrooke [15], who used it in his approach to determining optimal numbers of spares for a two echelon supply system.

4.3.2. A Repair Type as Solely a Supply Function

We may also select one of the repair types to represent new procurement. This would correspond to catastrophic failure, with repair consisting of buying a new unit. In this case, expected repair time consists of the expected time to order, receive, and transport a new unit to the pool of spares.

4.3.3. One For One Replenishment Required

In any case where one of the repair types is interpreted as having a supply function, we are making the assumption that a (continous review) $(s-1, s)$ policy is being followed, rather than, for example an (s, S) policy, and batch orders are not allowed.

4.4. Failures Generated by a Poisson Process

We assumed that, for each item at each location, $\{N(t);\ t \geqslant 0\}$ is a stationary Poisson process with rate $\lambda (\lambda > 0)$. That is, the time between failure follows a negative exponential distribution with mean $1/\lambda$, which is independent of time and the number of operational units. In particular, λ can depend on n, the *maximum* number of operational units, but not on the *actual* number, which may be less than n.

This assumption contrasts with a common assumption that the mean time between failures does depend on the number of operational units. This latter situation occurs when, for example, a unit can fail only if it is operating (installed as one of the operational units) and if, given that a unit is operating, the time until it fails is a random variable which (a) is independent of the time until failure for all other operational units and (b) follows a negative exponential distribution.

However, there are several ways of giving plausibility to our assumption. The first is to assume that whenever a shortage occurs, an emergency "loan" is effected, whereby an operative spare is immediately made available to eliminate the shortage, and when an operative spare is delivered to the pool, it is returned to the emergency borrower. With this assumption, it is irrelevant whether or not the mean time between failures depends on the number of operational units, since this number would always equal n. With this interpretation, our approach is to minimize the expected number of units "on loan" subject to a budget constraint.

A second way is to assert that the mean time between failures does not depend on the number of operational units, but on, say, the "level of activity" of the system. For example, if the operational

units are installed on aircraft, and the mean time between failures for this particular item correlates better with the number of flying hours of the fleet, rather than the number of operationally ready aircraft, then our approach gains plausibility. This interpretation is still approximate, because when there are no operational units, there should be no failures. In practice, this circumstance hopefully happens very rarely, in which case, this effect should be negligible.

4.5. The Cost Functions

4.5.1. Equivalent Stationary Cost

Recall that for a given item at a given location, $c^S(s)$ denotes the cost of retaining s spares in the system and $c_k^R(T_k)$ the cost of specifying T_k as the expected type k repair time, for each k. One reasonable way of determining these costs, in practice, is to use what we call equivalent stationary cost, also called equivalent average cost (e.g., [18]).

By equivalent stationary cost of an activity, we mean that amount, which spent evenly over the infinite horizon $[0, \infty)$ is equivalent to the actual expenditures incurred by that activity over $[0, \infty)$. For example, if expenditures occur periodically and the cost of capital is r per dollar per period, and c_n denotes the expenditure incurred at the beginning of period n for $n=0, 1, \ldots$, then the present value is $v \equiv \sum_{n=0}^{\infty} \alpha^n c_n$, where $\alpha = 1/(1+r)$, and the equivalent stationary cost is $v(1-\alpha) = rv/(1+r)$. Thus if the stationary cost $v(1-\alpha)$ were expended at the beginning of period n, for each n, the present value of this cost stream would also be v.

The appropriate selection of equivalent stationary cost when there is no cost of capital is the Césaro mean $\lim_{n\to\infty} \sum_{i=0}^{n} c_n/n$ if it exists.

Consider $c^S(s)$. At first glance, one might suggest that it be zero. This would occur if we assume (i) the cost of retaining s spares consists of merely initial procurement of the spares and (ii) there is no cost of capital. Taking into account new procurement required by irreparable failures would only add a fixed amount, which would not depend on s, to $c^S(s)$. Thus, adding spares would appear to have no additional cost. However, in practical economic problems, there is always a nonzero cost of capital. In addition, obsolescence would require periodic procurement of s units of the replacement item, and thus yields a cost which depends on s. Other costs, such as required storage space (particularly when the item is bulky and the location is a ship) would also be incorporated.

5. EVALUATION OF EXPECTED SHORTAGES

As discussed in section 3, we need only consider a fixed item and location in this section.

LEMMA 2: $S(s, \{T_k\})$ exists and equals $B(s, T) \equiv \sum_{x=s+1}^{\infty} (x-s)p(x; \lambda T)$,

where

$$p(x; \lambda T) \equiv \frac{e^{-\lambda T}(\lambda T)^x}{x!} \qquad \text{for } x=0, 1, \ldots,$$

$$T \equiv \sum_k p_k T_k, \text{ the expected repair time,}$$

and λ is the rate of the (Poisson) failure process.

REMARK: This result is used by Sherbrooke [15]. $S(s, \{T_k\})$ depends only on s and T; hence. the new notation $B(s, T)$.

PROOF: Because the failure process $\{N^F(t);\ t \geqslant 0\}$ is a stationary Poisson process and repair time are i.i.d. random variables, the distribution function of $N^R(t)$, the number of units in repair at time $t \geqslant 0$, is Poisson (e.g., [14, p. 18]). As $t \to \infty$, the probability that there are x units in repair at time t, approaches

$$p(x;\ \lambda T) \equiv e^{-\lambda T}(\lambda T)^x/x!.$$

The result follows immediately on recalling the definition of $S(s, \{T_k\})$. Q.E.D.

6. APPROACH TO THE OPTIMALIZATION PROBLEM

Our approach to solving (2) is to decompose it into two optimization problems. The first is: find positive numbers $T_1, T_2, \ldots, T_K$ to

$$\min \sum_k c_k^R(T_k) \tag{3.1}$$

$$\text{s.t.} \sum_k p_k T_k = T. \tag{3.2}$$

Define $c^R(T)$ to be the optimal value of (3.1), for the value of T specified in (3.2). We want to be able to solve (3) for each T (and for each item and location). We do this in section 8 for two special cases in which $c^R(T)$ can be written down in a simple closed form expression.

The second optimization problem is: find s and T to

$$\min L(s, T, u) \tag{4}$$

where

$$L(s, T, u) \equiv B(s, T) + uc^S(s) + uc^R(T).$$

The advantage of (4) over (2) is that here there are only two variables (for each item and location) rather than the $K+1$ which (2) has.

We now show that by solving (3) and (4), we solve (2).

LEMMA 3: If for fixed u, s^* and T^* solve (4) and $T_1^*, \ldots, T_K^*$ solves (3) with T^* specified in (3.2), then $s^*, T_1^*, \ldots, T_K^*$ solves (2).

PROOF: Pick arbitrary s and $T_1, \ldots, T_K$ ($s \epsilon \{0, 1, \ldots\}$ and $0 < T_k < \infty$ for each k). Let $T \equiv \sum_k p_k T_k$. Then

$$\begin{aligned} S(s^*, T_1^*, \ldots, T_K^*) &= B(s^*, T^*) && [\text{Lemma 2 and } T^* = \textstyle\sum_k p_k T_k^*] \\ &\leqslant B(s, T) && [(s^*, T^*) \text{ solves (4)}] \\ &= S(s, T_1, \ldots, T_K) && [\text{Lemma 2}] \end{aligned}$$

which completes the result. Q.E.D.

7. EVALUATION OF $c^R(T)$

We supply two examples of cases where $c^R(T)$ can be written down in closed form; in fact, in either the form $a/(T-d)$ or $ae^{-(T-d)/r}$.

7.1. Case One: $c_k(T_k) = c_k/(T_k - d_k)$ for each k

Assuming c_k is positive, d_k is nonnegative, and $T_k > d_k$ for each k, we can write $c^R(T)$ in closed form (for $T > \sum_k p_k d_k$) in this case as follows. We form the Lagrangian (for fixed T) for (3):

$$L(T_1, T_2, \ldots, T_K, u) \equiv \sum_k c_k \Big/ (T_k - d_k) + u \left[\sum_k p_k T_k - T\right]$$

and calculate

$$\frac{\partial}{\partial T_k} L(T_1, \ldots, T_K, u) = -\frac{c_k}{(T_k - d_k)^2} + u p_k \tag{5}$$

and

$$\frac{\partial^2}{\partial T_k^2} L(T_1, \ldots, T_K, u) = \frac{2c_k}{(T_k - d_k)^3},$$

which is positive for $T_k > d_k$. Thus, since the cross partials vanish, by setting (5) equal to zero for each k, yielding

$$T_k^*(u) = d_k + \sqrt{\frac{c_k}{u p_k}}, \tag{6}$$

we have, for each fixed positive u, minimized L. Then, for fixed u, we have

$$\sum_k p_k T_k^*(u) = u^{-1/2} \sum_k (p_k c_k)^{1/2} + \sum_k p_k d_k$$

which, when set equal to T, yields

$$u^* = a/(T-d)^2,$$

where

$$a \equiv \left[\sum_k (p_k c_k)^{1/2}\right]^2$$

and

$$d \equiv \sum_k p_k d_k.$$

Thus, substituting into (6), we have

$$T_k^* \equiv T_k^*(u^*) = d_k + (T-d)[c_k/(a p_k)]^{1/2}$$

so that

$$c^R(T) \equiv \sum_k c_k/(T_k - d_k) = a/(T-d).$$

That is, if the cost functions for the expected repair times for each type of repair are of the form $a/(T-d)$, then the cost function for the (overall) expected repair time is of the same form.

7.2. Case Two: $c_k(T_k) = c_k e^{-(T_k - d_k)/r_k}$ for each k

Assuming here that c_k and r_k are positive, d_k is nonnegative, and $T_k > d_k$ is required, we can again write $c^R(T)$ in closed form, for $T > d \equiv \sum_k p_k d_k$. This result will hold only in certain cases, however, as indicated below.

Here

$$\frac{\partial}{\partial T_k} L(T_1, \ldots, T_k, u) = -(c_k/r_k)e^{-(T_k - d_k)/r_k} + up_k$$

which yields

$$T_k^*(u) = d_k + r_k \ln\left(\frac{c_k}{up_k r_k}\right)$$

so that

$$\sum_k c_k(T_k^*(u)) = ur, \tag{7}$$

where

$$r \equiv \sum_k p_k r_k.$$

Picking u so that

$$\sum_k p_k T_k^*(u) = T$$

yields

$$u^* = (a/r)e^{-(T-d)/r},$$

where

$$a \equiv r \exp\left\{\left[\sum_k p_k r_k \ln\left(\frac{c_k}{p_k r_k}\right)\right] \Big/ r\right\}.$$

Thus, on assuming $T_k^*(u^*) \geqslant d_k$ for each k and then substituting into (7), we have

$$c^R(T) = ae^{-(T-d)/r},$$

so that the form of the cost function is again preserved.

8. MINIMIZING $L(s, T, u)$ For Fixed u

8.1. Preliminaries

LEMMA 4. $B(s, T)$ is

1° for fixed T, strictly decreasing and discretely convex in s, and

2° for fixed s, continuously differentiable, strictly increasing, and strictly convex in T.

PROOF: Recall that $p(x; \lambda T) \equiv e^{-\lambda T}(\lambda T)^x/x!$ for $x = 0, 1, \ldots$. Let, for $s = 1, 2, \ldots$,

$$\Delta_1 B(s, T) \equiv B(s, T) - B(s-1, T). \tag{8}$$

Then, by direct computation

$$\Delta_1 B(s,T) = -\sum_{x=s}^{\infty} p(x;\lambda T) < 0, \text{ and}$$

$$\Delta_{11} B(s,T) \equiv \Delta_1(\Delta_1 B(s,T)) = p(s-1;\lambda T) > 0,$$

from which 1° follows.

Again, by direct computation, we see that

(10) $$\frac{\partial}{\partial T} p(x;\lambda T) = \lambda[p(x-1;\lambda T) - p(x;\lambda T)]$$

so that

(11) $$\frac{\partial B(s,T)}{\partial T} = \lambda \sum_{x=s}^{\infty} p(s;\lambda T) > 0$$

and

$$\frac{\partial^2}{\partial T^2} B(s,t) = \lambda^2 p(s-1;\lambda T) > 0,$$

from which 2° follows Q.E.D.

8.2. Coordinate Direction Minimization

Assume in the sequel that

1* $c^S(s)$ is (discretely) convex in s, and

2* $c^R(T)$ is continuously differentiable and convex in T.

Also assume, unless indicated otherwise, that $s \epsilon \{0, 1, \ldots\}$ and $T, u \epsilon (0, \infty)$.

LEMMA 5:

1° $L(s,T,u)$ is, for fixed T and u, (discretely) convex in s,

2° $L(s,T,u)$ is, for fixed s and u, continuously (partially) differentiable and strictly convex in T,

3° $\frac{\partial}{\partial T} L(s,T,u)$ is, for fixed T and u, strictly decreasing in s, and

4° $\Delta_1 L(s,T,u)$ is, for fixed s and u, strictly decreasing in T.

PROOF: 1° and 2° follow directly from 1*, 2*, and Lemma 4, since,

(12) $$L(s,T,u) \equiv B(s,T) + uc^S(s) + uc^R(T),$$

and $u \geqslant 0$.

3° and 4°: By (9), (10), (11), (12), and direct calculation,

$$\Delta_1 \frac{\partial}{\partial T} L(s,T,u) = \frac{\partial}{\partial T} \Delta_1 L(s,T,u) = -\lambda p(s-1;\lambda T) \qquad \text{if } s \geqslant 1,$$

which is strictly negative. Q.E.D.

Let $s^*(T, u)$ denote the set of nonnegative integers which minimize $L(\cdot, T, u)$ for fixed T and u, and $T^*(s, u)$ the set of positive real numbers which minimize $L(s, \cdot, u)$ for fixed s and u. For example, if $s^\circ \epsilon s^*(T, u)$, then

$$L(s^\circ, T, u) \leqslant L(s, T, u) \qquad \text{for all } s.$$

Assume hereafter, for convenience, that $s^*(T, u)$ and $T^*(s, u)$ are nonempty.

LEMMA 6: Consider

1° $s \epsilon s^*(T, u)$ and $s \geqslant 1$,

2°

$$\Delta_1 L(s, T, u) \leqslant 0 \tag{13}$$

and

$$\Delta_1 L(s+1, T, u) \geqslant 0, \tag{14}$$

3° $0 \epsilon s^*(T, u)$,

4° (14) holds for $s = 0$,

5° $T \epsilon T^*(s, u)$, and

6°

$$\frac{\partial}{\partial T} L(s, T, u) = 0. \tag{15}$$

Then 1° iff 2°; 3° iff 4°; and 5° iff 6°.

REMARK: Thus, by 2° of Lemma 5, $T^*(s, u)$ consists of one element, for each s and u, which, for convenience, we also denote by $T^*(s, u)$.

PROOF: 1° iff 2°: (13) and (14) are clearly necessary conditions. They are sufficient, by 1° of Lemma 5.

3° iff 4°: Here, $\Delta_1 L(0, T, u)$ is undefined, prompting the special case, otherwise the same as above.

5° iff 6°: (15) is clearly necessary. It is sufficient, by 2° of Lemma 5. Q.E.D.

We say (s°, T°) is *u-optimal* if $L(s^\circ, T^\circ, u) \leqslant L(s, T, u)$ for all s and T; i.e., (s°, T°) minimizes $L(\cdot, \cdot, u)$ for fixed u.

LEMMA 7: (Necessary conditions). If

1° (s, T) is u-optimal, then

2° $s \epsilon s^*(T, u)$ and $T \epsilon T^*(s, u)$.

PROOF: This follows by contradiction. For example, if $s \notin s^*(T, u)$, there is some $s^\circ \epsilon s^*(T, u)$ such that $L(s^\circ, T, u) < L(s, T, u)$ which contradicts 1°. Q.E.D.

Recall that our approach starts with a sequence $\{u_k\}$ of multipliers and then finds an associated sequence of undominated solutions $\{(s^i, T^i)\}$ (where (s^i, T^i) is u_i-optimal for each i). Thus, (13)–(15) must be satisfied by such solutions. However, it is important to note that (13)–(15) are not sufficient conditions. Similar difficulties in a related model were pointed out by Veinott [17]. Miller [11] defines "discrete convexity" so that the conditions are sufficient when $L(\cdot,\cdot,u,)$ is (discretely) convex for fixed

u. However, his conditions are not satisfied in our case. In fact, we can find examples for our problem where $s \epsilon s^*(T, u)$ and $T \epsilon T^*(s, u)$ (so that (s, T) is a fixed point of the cyclic coordinate ascent method (e.g., [19]), yet (s, T) does *not* minimize $L(\cdot, \cdot, u)$ (even over a small neighborhood of (s, T)).

We now outline a partial enumeration algorithm which will minimize $L(\cdot, \cdot, u)$ for fixed u, under the assumption that there exist lower and upper bounds, say $\underline{s}$ and $\bar{s}$, on s^*.

1° set $s = \underline{s}$.

2° Find $T^*(s, u)$ and set $T = T^*(s, u)$.

3° Store $L(s, T, u)$ if it is the smallest yet found.

4° If $s = \bar{s}$, stop.

5° Replace s by $s + 1$ and repeat from 2°.

Any of several one dimensional search schemes (e.g., [18]) may be used in step 2°. It is clear that, given the lower and upper limits on feasible values of s, the algorithm will, for fixed u, find a u-optimal solution (subject to the constraints on s). In the next section, for specific cases, we show how to derive bounds not only on s but on T for any u-optimal solution, thereby allowing the algorithm to be used with no loss in optimality.

Before doing so, however, we indicate a result which yields bounds which hold in general.

LEMMA 8: If

1° $$s_1 < s_2, T_1 \epsilon T^*(s_1, u), \qquad \text{and } T_2 \epsilon T^*(s_2, u),$$

then

2° $$T_1 < T_2.$$

Similarly, if

3° $$T_1 < T_2, s_1 \epsilon s^*(T_1, u), \qquad \text{and } s_2 \epsilon s^*(T_2, u),$$

then

4° $$s_1 \leqslant s_2.$$

REMARK: That 1° implies 2° is useful in step 2° of our algorithm. That is, as the number of spares is increased, the "optimal" (for that number of spares) expected repair time increases. This allows iterative improvement of the lower bound on the "optimal" expected repair times for successive trial values of s.

PROOF: 1° implies 2°: Suppose $T_2 \leqslant T_1$. Then

$$0 = \frac{\partial}{\partial T} L(s_2, T_2, u) \qquad [(15)]$$

$$< \frac{\partial}{\partial T} L(s_1, T_2, u) \qquad [3° \text{ of Lemma 5}]$$

$$\leqslant \frac{\partial}{\partial T} L(s_1, T_1, u) \qquad [2° \text{ of Lemma 5}]$$

$$= 0 \qquad [(15)]$$

a contradiction.

3° implies 4°: Similarly, suppose $s_1 > s_2$ so that $s_1 \geqslant s_2 + 1 \geqslant 1$. Then,

$$\begin{aligned} 0 &\leqslant \Delta_1 L(s_2+1, T_2, u) && [(14)] \\ &\leqslant \Delta_1 L(s_2+1, T_1, u) && [4° \text{ of Lemma 5}] \\ &\leqslant \Delta_1 L(s_1, T_1, u) && [1° \text{ of Lemma 5 and } s_1 \geqslant s_2+1] \\ &\leqslant 0 && [(13)] \end{aligned}$$

a contradiction. Q.E.D.

9. QUALITATIVE RESULTS, SPECIAL CASES

In the next two subsections, we examine two special cases, under the assumption

$$1^{**} c^S(s) = c \cdot s,$$

where $c > 0$.

For this assumption, we have the following.

LEMMA 9: Suppose (s, T) is u-optimal. Then

1°

(16) $$\lambda \sum_{x=s}^{\infty} p(x; \lambda T) + u \frac{\partial}{\partial T} c^R(T) = 0,$$

2° if $uc \geqslant 1$, then $s = 0$, and

3° if $uc < 1$, then

(17) $$\sum_{x=s}^{\infty} p(x; \lambda T) \geqslant uc,$$

and

(18) $$\sum_{x=s+1}^{\infty} p(x; \lambda T) \leqslant uc.$$

PROOF: (16) is simply (15), which must hold by Lemmas 6 and 7.

2°: By 1**, (14) becomes (18), which holds, since

$$\sum_{x=s+1}^{\infty} p(x: \lambda T) < 1 \leqslant uc.$$

That is, $L(0, T, u) < L(1, T, u)$.

3°: (13) and (14) have become (17) and (18) here, which, by Lemmas 6 and 7, must hold if $s \geq 1$. If $s=0$, then (18) must hold by Lemma 6, and (17) holds, since $\sum_{x=0}^{\infty} p(x; \lambda T) = 1 > uc$. Q.E.D.

9.1. Case One: $c^R(T) = a/(T-d)$

Recall from section 7.1 the circumstances under which this assumption is justified. Here

$$L(s, T, u) = B(s, t) + ucs + ua/(T-d),$$

and (16) becomes

$$\sum_{x=s}^{\infty} p(x; \lambda T) = ua/[\lambda (T-d)^2]. \tag{19}$$

Theorem 1: In this case, if (s, T) is u-optimal, then

1° if $uc \geq 1$, then $s=0$ and $T=\underline{T}$,

2° if $uc < 1$ then

where $\underline{T} \leq T \leq \bar{T}$ and $\underline{s} \leq s \leq \bar{s}$,

$$\bar{T} \equiv d + \sqrt{a/\lambda c}, \tag{20}$$

$$\underline{T} \equiv d + \sqrt{ua/\lambda}, \tag{21}$$

$\bar{s}$ is the smallest integer which satisfies

$$\sum_{x=\bar{s}+1}^{\infty} p(x; \lambda \bar{T}) < uc, \tag{22}$$

and $\underline{s}$ is the largest integer which satisfies

$$\sum_{x=\underline{s}}^{\infty} p(x; \lambda \underline{T}) > uc. \tag{23}$$

REMARK: These bounds allow the algorithm of section 8.2 to be used, with no loss in optimality.

PROOF:

1°: By Lemma 9, $s=0$, $T=\underline{T}$ follows from (19) and (21).

2°: $T \leq \bar{T}$: Here.

$$ua/[\lambda (T-d)^2] = \sum_{x=s}^{\infty} p(x; \lambda T) \qquad [(19)]$$

$$\geq uc \qquad [(17)]$$

from which the result follows.

$\underline{T} \leqslant T$: Here,

$$ua/[\lambda(T-d)^2] = \sum_{x=s}^{\infty} p(x;\lambda T) \leqslant 1$$

which yields the result.

$s \leqslant \bar{s}$: Suppose $s > \bar{s}$, which implies $s \geqslant \bar{s}+1 \geqslant 1$. First observe that, by (10), since $s \geqslant 1$,

$$\frac{\partial}{\partial T} \sum_{x=s}^{\infty} p(x;\lambda T) = \lambda p(s-1;\lambda T) > 0. \tag{24}$$

Then

$$\begin{aligned} uc &\leqslant \sum_{x=s}^{\infty} p(x;\lambda T) && [(17)] \\ &\leqslant \sum_{x=s}^{\infty} p(x;\lambda \bar{T}) && [(24)] \text{ and } T \leqslant \bar{T}] \\ &\leqslant \sum_{x=\bar{s}+1}^{\infty} p(x;\lambda \bar{T}) && [s \geqslant \bar{s}+1] \\ &< uc && [(22)], \end{aligned}$$

a contradiction.

$\underline{s} \leqslant s$: Similarly, suppose $\underline{s} > s$, so that $\underline{s} \geqslant s+1 \geqslant 1$. Then

$$\begin{aligned} uc &\geqslant \sum_{x=s+1}^{\infty} p(x;\lambda T) && [(18)] \\ &\geqslant \sum_{x=s+1}^{\infty} p(x;\lambda \underline{T}) && [(24), s+1 \geqslant 1, \text{ and } \underline{T} \leqslant T] \\ &\geqslant \sum_{x=\underline{s}}^{\infty} p(x;\lambda \underline{T}) && [\underline{s} \geqslant s+1] \\ &> uc, && [(23)] \end{aligned}$$

a contradiction. Q.E.D.

9.2. Case Two: $c^R(T) = ae^{-(T-d)/r}$

Recall from section 7.2 the justification for this case. Here

$$L(s,T,u) = B(s,T) + ucs + uae^{-(T-d)/r},$$

and (16) becomes

$$\sum_{x=\underline{s}}^{\infty} p(x; \lambda T) = (uae^{-(T-d)/r})/(r\lambda). \tag{25}$$

THEOREM 2: In this case, if (s, T) is u-optimal, then

1° If $uc \geqslant 1$, then $s = 0$ and $T = T$, and

2° If $uc < 1$, then $\underline{T} \leqslant T < \bar{T}$ and $\underline{s} \leqslant s \leqslant \bar{s}$, where

and

$$\bar{T} = d + r \ln (a/(r\lambda c)), \tag{26}$$

$$T \equiv d + r \ln (ua/(r\lambda)) = r \ln (uc) + \bar{T}, \tag{27}$$

and $\underline{s}$ and $\bar{s}$ are given as in (22) and (23).

REMARK: For convenience, we use the same notation for our bounds, although their values differ from those of case one.

PROOF: The proof follows that of Theorem 1, and is therefore omitted. Q.E.D.

10. A NUMERICAL EXAMPLE

Next we illustrate our results by considering a two item, two location problem. Suppose the locations are different ships. Let s_1 denote the number of spares for item A on board ship #1, s_2 the number of spares of item A on ship #2, and s_3 the number of spares of item B on ship #2. (For convenience, we assume ship #1 does not carry item B). Here, T_{ik} is the mean repair time for the kth echelon of repair for the ith item. In this example there are three types of repair: $K=1$ refers to local repair, $k=2$ refers to repair at another facility, and $k=3$ refers to resupply on a one-for-one basis. We assume in this example that the cost functions satisfy $c_i^S(s_i) = c_i^S \cdot s_i$ and $c_{ik}^R(T_{ik}) = c_{ik}^R / T_{ik}$. Thus, our problem is to determine s_i and T_{ik} to

$$\min \sum_{i=1}^{3} B(s_i, T_i) \tag{28.1}$$

i.e., expected total backorders over both items and both ships subject to

$$T_i = \sum_{k=1}^{3} T_{ik} p_{ik}, \text{ and} \tag{28.2}$$

$$\sum_{i=1}^{3} \left[c_i^S \cdot s_i + \sum_{k=1}^{3} c_{ik}^R / T_{ik} \right] \leqslant b, \tag{28.3}$$

where $B(s_i, T_i)$ is the expected number of shortages for item i, b is the available budget, and p_{ik} is the probability that the kth type of repair will be needed when a unit of item i fails. Note that in this problem there are 12 decision variables. The values used for the parameters p_{ik}, c_{ik}^R, c_i^S, and λ_i are given in

Table 1. As discussed in section 3, our solution procedure automatically constructs the solutions for a sequence of budgets b. However, for the sake of illustration we will consider specifically the case $b = 115$.

Interpretation of subscript	p_{ik}			c_{ik}^R			c_i^S	λ_i
	$k=1$	$k=2$	$k=3$	$k=1$	$k=2$	$k=3$		
Item A, Ship #1......	0.20	0.40	0.40	20	40	40	1	5
Item A, Ship #2......	0.50	0.25	0.25	3.2	6.4	57.6	2	5
Item B, Ship #2......	0.50	0.25	0.25	24	48	432	3	1

Rather than solving (28) directly, section 3 shows how a Lagrangian approach can be used which allows the problem to be decomposed by item and location. Then section 6 shows how this latter problem can be furthur decomposed into two optimization problems: the first problem computes (for $i = 1, 2, 3$)

$$c_i^R(T_i) = \min \sum_{k=1}^{3} c_{ik}^R / T_{ik} \tag{29}$$

subject to (28.2); and the second problem determines s_i and T_i (for $i = 1, 2, 3$) to minimize the Lagrangian function

$$L_i(s_i, T_i, u) = B(s_i, T_i) + u \cdot c_i^S \cdot s_i + u \cdot c_i^R(T_i) \tag{30}$$

for different values of the multiplier u.

In this example the repair cost functions satisfy $c_{ik}^R(T_{ik}) = c_{ik}^R / T_{ik}$. Section 7.1 shows how to solve problem (29) with these functions and demonstrates that $c_i^R(T_i) = a_i / T_i$, where

$$a_i = \left[\sum_{k=1}^{3} (p_{ik} c_{ik}^R)^{1/2} \right]^2. \tag{31}$$

The parameter values in Table 1 imply that using (31). $a_1 = 100$, $a_2 = 40$, and $a_3 = 300$. Hence, $c_1(T) = 100/T$, $c_2(T) = 40/T$, and $c_3(T) = 300/T$.

Theorem 1 of section 9.1 provides the conditions of optimality for minimizing the Lagrangian function (30). Note that the conditions in this theorem uniquely specify the u-optimal solution for the case $u \cdot c_i^S \geqslant 1$, but only provide upper and lower bounds on s_i and T_i for the case $u \cdot c_i^S < 1$. In this latter case, we use the search algorithm given in section 8.2 to compute the u-optimal solutions. Typical results obtained with this algorithm are discussed next.

Tables 2–4 provide the u-optimal values of s_i and T_i, as a function of u, for each item. Table 2 considers item A on ship #1, where $\lambda = 5$, $a_1 = 100$, and $c_1^S = 1$; Table 3 considers item A on ship #2,

where $\lambda_2=5$, $a_2=40$, and $c_2^S=2$*; and Table 4 considers item B on ship #2, where $\lambda_3=1$, $a_3=300$, and $c_3^S=3$. In each of these tables, the first column provides the multiplier u; the second column gives the corresponding minimum value of the Lagrangian (30); the third and fourth columns give the u-optimal solutions s_i and T_i, respectively; the fifth column computes the cost of the solution, $c_i^S \cdot s_i + a_i/T_i$; the sixth column gives the expected number of backorder $B(s_i, T_i)$; and the remaining columns provide the upper and lower bounds on s_i and T_i which are implied by Theorem 1. For the cases in which the upper and lower bounds do not coincide, it was necessary to use the search algorithm of section 8.2 to compute the u-optimal solutions given in the third and fourth column.

TABLE 2. *Solution of the Lagrangian Problem for Item A on Ship #1 i.e., an Item with* $\lambda_1=5$, $a_1=100$, *and* $c_1^S=1$

u	$L_1(s_1, T_1, u)$	s_1	T_1	$c_1^S \cdot s_1 + a_1/T_1$	$B(s_1, T_1)$	Bounds on s_1		Bounds on T_1	
						Low	High	Low	High
1.00	44.72	0	4.47	22.36	22.36	0	0	4.47	4.47
0.96	43.31	14	4.44	36.53	8.24	14	14	4.38	4.47
0.92	41.81	16	4.45	38.48	6.41	15	16	4.29	4.47
0.88	40.26	17	4.44	39.51	5.49	16	17	4.20	4.47
0.84	38.67	17	4.38	39.86	5.19	16	18	4.10	4.47
0.80	37.05	18	4.39	40.76	4.44	16	18	4.00	4.47
0.76	35.41	19	4.43	41.59	3.80	16	19	3.90	4.47
0.72	33.74	19	4.36	41.91	3.56	16	19	3.80	4.47
0.68	32.06	19	4.30	42.25	3.33	16	20	3.69	4.47
0.64	30.35	20	4.35	42.98	2.84	16	21	3.58	4.47
0.60	28.62	20	4.29	43.31	2.64	16	21	3.46	4.47
0.56	26.88	21	4.35	43.99	2.24	16	21	3.35	4.47
0.52	25.11	21	4.29	44.33	2.06	16	22	3.23	4.47
0.48	23.33	21	4.22	44.69	1.88	16	22	3.10	4.47
0.44	21.52	22	4.29	45.34	1.58	15	23	2.97	4.47
0.40	19.70	22	4.22	45.71	1.42	15	23	2.83	4.47
0.36	17.86	23	4.28	46.36	1.17	15	24	2.68	4.47
0.32	16.00	23	4.21	46.78	1.03	14	24	2.53	4.47
0.28	14.12	24	4.26	47.46	0.83	14	25	2.37	4.47
0.24	12.21	24	4.18	47.94	0.70	13	26	2.19	4.47
0.20	10.28	25	4.22	48.70	0.54	13	26	2.00	4.47
0.16	8.32	25	4.11	49.32	0.43	12	27	1.79	4.47
0.12	6.33	26	4.12	50.26	0.30	11	28	1.55	4.47
0.08	4.30	27	4.10	51.42	0.18	10	29	1.27	4.47
0.04	2.21	28	3.98	53.11	0.08	8	31	0.89	4.47

Tables 2–4 can be interpreted in the following way. Each of these tables provides the solution of the single item, single ship problem with a budget constraint limited to that item. For example, consider the first row in Table 2. Here the multiplier $u=1$ is associated with the cost 22.36, According to Lemmas 1 and 3, the solution $s_1=0$ and $T_1=4.7$ would be optimal if the repair budget for that item were equal to 22.36. The other rows in these tables may be interpreted in the same way. Note that as the multiplier decreases (budget increases), the optimal number of spares increases (i.e., does not decrease). This is plausible, yet we were not able to prove that it must always be true. The optimal expected repair time usually decreases, but it does not always do so, as can be seen in Table 2.

*NOTE: To illustrate the generality of the approach, the marginal cost of providing spares of item A was assumed to be different on ship 1 than on ship 2. In practice, the marginal cost would most likely be identical.

TABLE 3. *Solution of the Lagrangian Problem for Item A on Ship #2 i.e., an Item with* $\lambda_2 = 5$, $a_2 = 40$, *and* $c_2^S = 2$

u	$L_2(s_2, T_2, u)$	s_2	T_2	$c_2^S \cdot s_2 + a_2/T_2$	$B(s_2, T_2)$	Bounds on s_2		Bounds on T_2	
						Low	High	Low	High
1.00	28.28	0	2.83	14.14	14.14	0	0	2.83	2.83
0.96	27.71	0	2.77	14.43	13.86	0	0	2.77	2.77
0.92	27.13	0	2.71	14.74	13.57	0	0	2.71	2.71
0.88	26.53	0	2.65	15.08	13.27	0	0	2.65	2.65
0.84	25.92	0	2.59	15.43	12.96	0	0	2.59	2.59
0.80	25.30	0	2.53	15.81	12.65	0	0	2.53	2.53
0.76	24.66	0	2.47	16.22	12.33	0	0	2.47	2.47
0.72	24.00	0	2.40	16.67	12.00	0	0	2.40	2.40
0.68	23.32	0	2.33	17.15	11.66	0	0	2.33	2.33
0.64	22.63	0	2.26	17.68	11.31	0	0	2.26	2.26
0.60	21.91	0	2.19	18.26	10.95	0	0	2.19	2.19
0.56	21.17	0	2.12	18.90	10.58	0	0	2.12	2.12
0.52	20.40	0	2.04	19.61	10.20	0	0	2.04	2.04
0.48	19.44	5	1.99	30.10	4.99	5	5	1.96	2.00
0.44	18.18	6	1.95	32.49	3.89	6	6	1.88	2.00
0.40	16.83	7	1.94	34.61	2.98	6	7	1.79	2.00
0.36	15.41	8	1.95	36.51	2.27	7	8	1.70	2.00
0.32	13.94	8	1.88	37.25	2.02	7	9	1.60	2.00
0.28	12.41	9	1.91	38.95	1.50	7	9	1.50	2.00
0.24	10.84	9	1.84	39.77	1.29	7	10	1.39	2.00
0.20	9.20	10	1.87	41.40	0.92	7	11	1.27	2.00
0.16	7.53	10	1.78	42.42	0.74	7	11	1.13	2.00
0.12	5.79	11	1.80	44.19	0.48	6	12	0.98	2.00
0.08	3.98	12	1.80	46.21	0.28	6	13	0.80	2.00
0.04	2.08	13	1.75	48.92	0.13	5	15	0.57	2.00

Next consider the original problem (28), which is a two-item, two-ship problem with a joint budget constraint. The total costs and expected shortages for this problem are equal to the sum of the costs and expected shortages for each individual item and ship. Thus Table 5 is constructed by summing the cost and backorder columns in Tables 2–4. Table 5 may be interpreted in the following way. Consider the first row in which the multiplier $u=1$ is associated with the cost 53.82 and expected shortages 53.82. It follows from Lemmas 1 and 3 that 53.82 is the minimum number of shortages in problem (28), subject to a budget $b=53.82$. The corresponding optimal values of s_i and T_i for each item are given in the first row of Tables 2–4. Similarly, the second row in Table 5 implies that 38.07 is the minimum expected number of shortages subject to a budget $b=68.64$. Thus Tables 2–5 provide a sequence of undominated solutions; i.e., optimal solutions to problems of the form (28) with different budget levels. Suppose the actual budget is $b=115$. Table 5 shows that the multiplier $u=0.36$ has a cost of 111.71, which is the largest cost less than 115 in this table. Thus Tables 2–3 imply that an approximate solution of problem (28) with $b=115$ is given by $s_1=23$, $T_1=4.28$, $s_2=8$, $T_2=1.95$, $s_3=0$, and $T_3=10.39$. In other words, there should be 23 spares of item A on ship #1, 8 spares of item A on ship #2, and no spares of item B on ship #2. In addition, the expected repair time for item A on ship #1, should be set at 4.28 days. Similarly, it should be 10.39 days for item B on ship #2. Improved approximate solutions can be obtained by repeating the algorithm with a sequence of multipliers in the range [0.32, 0.36].

TABLE 4. *Solution of the Lagrangian Problem for Item B on Ship #2 i.e., an Item with* $\lambda_3 = 1$, $a_3 = 300$, *and* $c_3^S = 3$

u	$L_3(s_3, T_3, u)$	s_3	T_3	$c_3^S \cdot s_3 + a_3/T_3$	$B(s_3, T_3)$	Bounds on s_3		Bounds on T_3	
						Low	High	Low	High
1.00	34.64	0	17.32	17.32	17.32	0	0	17.32	17.32
0.96	33.94	0	16.97	17.68	16.97	0	0	16.97	16.97
0.92	33.23	0	16.61	18.06	16.61	0	0	16.61	16.61
0.88	32.50	0	16.25	18.46	16.25	0	0	16.25	16.25
0.84	31.75	0	15.88	18.90	15.88	0	0	15.88	15.88
0.80	30.98	0	15.49	19.37	15.49	0	0	15.49	15.49
0.76	30.20	0	15.10	19.87	15.10	0	0	15.10	15.10
0.72	29.39	0	14.70	20.41	14.70	0	0	14.70	14.70
0.68	28.57	0	14.28	21.00	14.28	0	0	14.28	14.28
0.64	27.71	0	13.86	21.65	13.86	0	0	13.86	13.86
0.60	26.83	0	13.42	22.36	13.42	0	0	13.42	13.42
0.56	25.92	0	12.96	23.15	12.96	0	0	12.96	12.96
0.52	24.98	0	12.49	24.02	12.49	0	0	12.49	12.49
0.48	24.00	0	12.00	25.00	12.00	0	0	12.00	12.00
0.44	22.98	0	11.49	26.11	11.49	0	0	11.49	11.49
0.40	21.91	0	10.95	27.39	10.95	0	0	10.95	10.95
0.36	20.79	0	10.39	28.87	10.39	0	0	10.39	10.39
0.32	19.44	5	9.95	45.15	4.99	5	5	9.80	10.00
0.28	17.52	7	9.87	51.39	3.13	6	7	9.17	10.00
0.24	15.41	8	9.75	54.76	2.27	7	8	8.49	10.00
0.20	13.18	9	9.72	57.86	1.61	7	9	7.75	10.00
0.16	10.84	9	9.19	59.65	1.29	7	10	6.93	10.00
0.12	8.37	10	9.14	62.83	0.83	7	11	6.00	10.00
0.08	5.79	11	9.01	66.29	0.48	6	12	4.90	10.00
0.04	3.05	12	8.62	70.81	0.21	6	14	3.46	10.00

Once the optimal values of T_1, T_2, and T_3 have been obtained, the corresponding values for T_{ik} can be computed by using

$$T_{ik} = \frac{T_i (c_{ik}^R / p_{ik})^{1/2}}{\sum_{j=1}^{3} (p_{ij} c_{ij}^R)^{1/2}},$$

which was derived in section 7.1. We used this formula to compute T_{ik} from T_i for the case $b = 115$, using the values of c_{ik}^R and p_{ik} given in Table 1, and the resulting values for T_{ik} are given in Table 6. Note that this solution allocates the largest portion of the budget to the first item, as this is where the greatest reduction in shortages per dollar can be achieved. As indicated at the beginning of this paper, it is straightforward to minimize the weighted sum of expected shortages, instead of the unweighted sum (all weights are one) in (28.1), where the weights are input parameters. This algorithm can also easily handle a problem with much larger numbers of items, facilities, and repair types.

Note that for item A on ship #2, the local repair time should be 0.78 days, whereas the repair turnaround time at the next higher echelon (including transit to and from the ship) should be 1.56 days. This is much lower than that for item A on ship #1, or for item B on ship #2, since it was assumed the costs to lower the repair times for failing item A's on board ship #2 was much lower than for the other two cases (see Table 1).

TABLE 5. *Undominated Solutions of the Two Item, Two Ship—Three Echelon Repair System*

u	$\sum_i (c_i^S \cdot s_i + a_i/T_i)$	$\sum_i B(s_i, T_i)$
1.00	53.82	53.82
0.96	68.64	39.07
0.92	71.28	36.58
0.88	73.05	35.01
0.84	74.19	34.02
0.80	75.94	32.58
0.76	77.68	31.23
0.72	78.99	30.26
0.68	80.40	29.28
0.64	82.31	28.01
0.60	83.93	27.01
0.56	86.04	25.78
0.52	87.96	24.75
0.48	99.79	18.87
0.44	103.94	16.95
0.40	107.71	15.36
0.36	111.74	13.83
0.32	129.18	8.04
0.28	137.79	5.46
0.24	142.47	4.26
0.20	147.96	3.07
0.16	151.40	2.46
0.12	157.28	1.61
0.08	163.92	0.95
0.04	172.84	0.43

It is also of interest to note that this repair design and spare stock mix for a budget of 111.74 gives rise to expected shortages of 1.17 units for item A at ship #1, 12.27 units for item A at ship #2, and 10.39 units for item B at ship #2. Thus the total minimum expected shortages for items A and B at ships #1 and #2, for this budget level, is 13.83 units. Observe the striking decreases in expected shortages possible by modest increases in the budget. Increasing the budget from 111.74 to 129.18 (an increase of 16 percent) decreases shortages from 13.83 to 8.04 (a decrease of 42 percent). Increasing the budget by 55 percent decreases shortages by 97 percent.

TABLE 6. *Approximate Solutions for* b = 115

Item i	s_i	T_i	$k=1$	$k=2$	$k=3$	$B(s_i, T_i)$	$c_i^S \cdot s_i + a_i/T_i$
Item A, Ship 1	23	4.28	4.28	4.28	4.28	1.17	46.36
Item A, Ship 2	8	1.95	.78	1.56	4.68	2,27	36.51
Item B, Ship 2..................	0	10.39	4.16	8.31	24.94	10.39	28.87

11. CONCLUSION AND RELATED LITERATURE

We have constructed a normative model which considers the optimal design of a certain repair and supply system. The model seeks an optimal design of both the repair and supply sides of the system, under a common budget constraint. Other normative models of this type of system have tended to focus

on the supply side of the system. The model of Sherbrooke [15] is a good representative of these. In fact, our approach is closely related to his. The similarities are that his is a multi-item, multi-location model with several operational units; and it has several repair types (only two are mentioned, but the generalization is immediate), a general repair distribution (but repair times are i.i.d. random variables, as discussed in section 4.2), and control over the number of spares at each location. The differences are that the failure process is compound Poisson, there is control over the number of spares kept for one of the repair types (a higher echelon of supply), and there is no control over the repair times.

Models which consider the optimal design of queueing systems, such as in Hillier [9], can be thought to consider the repair side of the system. The only model we could find which focuses on both sides of the system is that of Jacobson [10], who applies marginal analysis and partial enumeration in his study of a multi-item, single location model with several operational units; a Poisson failure process; one repair type; identical exponential repair times for each item; a finite number of repair facilities; and control over the spares and the number of repair facilities.

Related models, which focus on characterization rather than optimization, are summarized by Barlow and Proschan [1, chap. 5]. In these models, failures occur only when a unit is operational. One model has one operational unit; general failure times; exponential repair times; and a finite number of repair facilities. A similar model, with exponential failure times, is discussed by Natarajan [12]. Another model, due to Takacs [16], has several operational units; exponential failure times; general repair times; and one repair facility.

An important obstacle to using our model will be that of estimating the cost functions corresponding to selecting various expected repair times for each type of repair and each item/location combination. We have assumed that these are separable, depending only on the type of repair and the item/location combination. As discussed in section 4.2, the changes possible on the repair side of the system are likely to consist of numbers and types of equipment and manning levels for various specialties, at the repair facilities. Drawing a correspondence between these changes and expected repair times is likely to be difficult, and, at best, approximate. However, even crude approximation may yield valuable insights into the tradeoffs possible between the repair and supply sides of the system.

ACKNOWLEDGEMENT

We wish to thank Professor Edward Silver for bringing the work of Demmy [21] to our attention, after our work here had been completed. Our work is similar to his, in that he addressed a similar problem and also took a Lagrangian approach.

REFERENCES

[1] Barlow, R. E. and F. Proschan, *Mathematical Theory of Reliability* (John Wiley and Sons, Inc., New York, 1965).

[2] Brooks, R. B. S. and A. M. Geoffrion, "Finding Everett's Lagrange Multipliers by Linear Programming," Operations Research *14*, 1149–1153 (1966).

[3] Dantzig, G. B., *Linear Programming and Extensions* (Princeton University Press, Princeton, 1963).

[4] Dantzig, G. B. and P. Wolfe, "Decomposition Principle for Linear Programs," Operations Research *8*, 101–111 (1960).

[5] Demmy, W. S., "Allocation of Spares and Repair Resources to a Multi-component System," AFLC Report 70–17, Wright-Patterson AFB, Ohio (1970).

[6] Everett, H., "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," Operations Research *11*, 399–417 (1963).

[7] Fox, B. L. and D. M. Landi, "Searching for the Multiplier in One-Constraint Optimization Problems," Operations Research *18*, 253–262 (1970).

[8] Gould, F. J., "Extensions of Lagrange Multipliers in Nonlinear Programming," SIAM J. Applied Mathematics *17*, 1280–1297 (1969).

[9] Greenberg, H. J., "Finding Everett's Lagrange Multipliers by Generalized Linear Programming," Technical Report CP–70008, Computer Science/Operations Research Center, Southern Methodist University, Dallas, Texas (1972).

[10] Hillier, F. S., "Economic Models for Industrial Waiting Line Problems," Management Science *10*, 119–130 (1963).

[11] Jacobson, L. J., "Optimum Allocation of Resource in the Purchase of Spare Parts and/or Additional Service Channels," ORC 69–12, University of California, Berkeley (1969).

[12] Miller, B. L., "Unconstrained Optimization in the Integers," RM–6165, The Rand Corporation, Santa Monica (1970).

[13] Natarajan, R., "A Reliability Problem with Spares and Multiple Repair Facilities," Operations Research *16*, 1041–1057 (1968).

[14] Parzen, E., *Stochastic Processes* (Holden-Day, San Francisco, 1962).

[15] Ross, S. M., *Applied Probability Models with Optimization Applications* (Holden-Day, San Francisco, 1970).

[16] Sherbrooke, C. C., "METRIC: A Multi-Echelon Technique for Recoverable Item Control," Operations Research *16*, 122–141 (1968).

[17] Takacs, L., "On a Combined Waiting Time and Loss Problem Concerning Telephone Traffic," *Ann. Univ. Scient. Budapest*, Eotvos, Sect. Math. 1 (1958), pp. 73–82.

[18] Veinott, A. F., Jr., "Review of *Analysis of Inventory Systems*, by G. Hadley and T. M. Whitin," J. Am. Stat. Assoc. *59*, 283–285 (1964).

[19] Wagner, H. M., *Principles of Operations Research* (Prentice-Hall, Englewood Cliffs, New Jersey, 1969).

[20] Zadeh, N., "A Note on the Cyclic Coordinate Ascent Method," Management Science *16*, 642–644 (1970).

[21] Zangwill, W. I., *Nonlinear Programming, A Unified Approach* (Prentice-Hall, Englewood Cliffs, 1969).

A GENERALIZED RECURSIVE ALGORITHM FOR A CLASS OF NON-STATIONARY REGENERATION (SCHEDULING) PROBLEMS

Fred Glover

Graduate School of Business
University of Colorado

and

Theodore D. Klastorin

Babcock Graduate School of Management
Wake Forest University

ABSTRACT

Recent efforts in the field of dynamic programming have explored the feasibility of solving certain classes of integer programming problems by recursive algorithms. Special recursive algorithms have been shown to be particularly effective for problems possessing a 0–1 attribute matrix displaying the "nesting property" studied by Ignall and Veinott in inventory theory and by Glover in network flows.

This paper extends the class of problem structures that has been shown amenable to recursive exploitation by providing an efficient dynamic programming approach for a general transportation scheduling problem. In particular, we provide alternative formulations tor the scheduling problem and show how the most general of these formulations can be readily solved vis a vis recursive techniques.

1. INTRODUCTION

In this paper, we develop scheduling models and associated solution methods for an aspatial transportation system. Our methods have application not only to such an aspatial system, but also to a large class of scheduling and allocation problems that exhibit similar structural characteristics.

We first show that the problem can be formulated in terms of decision variables representing the time between each action (decision event). This formulation is related to the infinite state regeneration model (with unbounded horizon) proposed by Wagner [1], particularized to case of a unit discount factor. However, our model is more flexible because it does not require the assumption of a constant revenue at each regeneration point. Furthermore, the revenue is permitted to be a function of either a deterministic or probabilistic demand.

The "natural" formulation of this model turns out to be extremely resistant to the development of efficient solution methods. We show that it is possible to provide an alternative formulation in terms of different decision variables which, in the transportation context, represent the number of vehicles to schedule at each period (considering each time period to be a regeneration point). Although this formulation is in fact a nonlinear integer programming model, we show that its structure can be efficiently exploited by a specially designed dynamic programming recursion. Moreover, the solution method we propose is shown to be applicable to more general models possessing a 0–1 attribute matrix

that satisfies a "nesting property" in its rows and columns, thus extending the domain of problem classes (and solution methods) for which such a property has been found relevant.*

Several antecedents of our work deserve mention. An excellent survey of airline scheduling models by Götz Uebe [5] discusses models due to Bisbee [6] and Tingaud [7] which are very similar to the prototype model of Klastorin [4] which provides the starting point for this paper. The procedures reported by Uebe can be straightforwardly adapted to Klastorin's initial model, but do not accommodate the more general problem structures examined here. (For other related dynamic programming models, see also Uebe [8].)

A recent paper on production planning with concave costs and capacity constraints by Michael Florian and Morton Klein [9], which usefully extends the fundamental papers of Zangwill [10], Wagner [11], and Wagner and Whitin [12], is still more closely related to our work. The model of the present paper generalizes the model of Florian and Klein to include nonconcave relationships in the objective function and additional nonlinearities in the constraints. The more complex structure of the model examined here does not admit the elegant extreme point characterizations of optimal solutions presented in Florian and Klein, but can nevertheless be accommodated, as we show, by a particularly efficient dynamic programming approach. In fact, whereas previous related recursions for handling nonconcavity (e.g., those of [11, 12]) involve multiple state variables which give rise to excessive amounts of computation, our approach involves only a single state variable, and hence for the class of models we examine is decidedly more convenient.

2. STATEMENT OF THE PROBLEM

2.1. Assumptions

The prototype for the class of problems discussed in this paper was first examined by Klastorin [4] in a study of aspatial transportation systems operating under decentralized control (i.e., a system which exhibits minimal communication among the terminals, transit, and control subsystems). The objective of the prototype model is to arrive at a profit maximizing schedule based only on demand and cost per vehicle trip. The time span under consideration (usually, but not necessarily, one day) is is broken into n discrete time intervals. Furthermore, in the prototype model, demand is assumed to be sufficiently known to be considered deterministic in each time period and the cost per vehicle trip is assumed to be constant (and positive) for each time period. The latter assumption implies that the number of passengers on any vehicle has no effect whatsoever on the cost of that vehicle trip. That is, a vehicle which is only half filled incurs the same expense as one which is completely full—provided, of course, that the routes are identical. Furthermore, all passengers pay the same fare for vehicles departing in a given period and taking a given route.

Under the assumptions of the model, the routes are operated with sufficient independence that interactions between routes have negligible influence on the determination of an optimal schedule. Hence, multiple routes may be accommodated simply by aggregating the results obtained from applying the model to each of these routes singly.

In what follows, we will adhere to the convention whereby vehicle trips scheduled "in" a given time period refer to those trips scheduled to leave at the end of that time period. Thus, all people

*See, e.g., Ignall and Veinott [2] for application in inventory theory and Glover [3] for applications in network flows.

entering the terminal subsystem during the period may be considered candidates for service by vehicles scheduled in that period. This assumption is of minimal concern; as the number of time periods gets sufficiently large, the number of passengers entering in any one time period becomes arbitrarily small.

2.2. A Direct Formulation

Mathematically, the problem can be stated in terms of decisions variables δ_{ij} representing the number of time periods (measured in minutes, hours, etc.) between the ith and jth vehicle departures. (If δ_{ij} equals 0, two or more vehicle departures are scheduled at the same time.) By convention, δ_{01} denotes the number of time periods between the start of the scheduling period and the first vehicle departure. Our definition of the decision variables here follows the standard format of the so-called *regeneration models* (see e.g., Wagner [1], pp. 375–378), and, indeed, the problem under consideration may be considered an instance of such models inasmuch as it directly concerns the optimization of regenerative phenomena (vehicle departures). In particular, in the limiting case, the present model strongly resembles the "infinite state regeneration model" of [1], except for certain nonlinearities that reflect the relationship between demand and the number of passengers accommodated.

Representing incremental demand (number of new passengers entering the terminal subsystem) in time period p by h_p and representing the capacity of each vehicle by γ, the total number of passengers T_k that are accommodated by the kth vehicle must satisfy:

$$T_k \leq \operatorname{Min}\left\{\gamma, \sum_{p=0}^{\theta_k} h_p - \sum_{q=1}^{k-1} T_q\right\} \qquad \text{for all } k=1, 2, \ldots, N$$

where N is the total number of vehicles scheduled to depart and θ_k is the number of time periods elapsing before the kth vehicle's departure; i.e.,

$$\theta_k = \sum_{i=0}^{k-1} \delta_{i,\,i+1}\,.$$

Letting n denote the total number of time periods under consideration, it follows that θ_k cannot exceed n for all values of k (hence, in particular, $\theta_N \leq n$). Finally to complete the notation, we let C_p denote the passenger fare and d_p denote the trip cost for vehicles departing in period p. Then the objective of maximizing total profit may be expressed as:

$$\text{Maximize} \sum_{k=1}^{N} (c_{\theta_k} T_k - d_{\theta_k}).$$

2.3 An Improved Formulation

Although the preceding formulation is both standard and "intuitive," the manner in which the decision variables δ_{ij} enter the objective function and the determination of the quantities T_k (through the numbers θ_k) produces an exceptionally difficult nonlinear structure that does not readily accommodate itself to the development of efficient solution methods. Consequently, we shall propose an alternative formulation which is more amenable to solution analysis and, in fact, can be extended to accommodate a variety of more general models for which efficient algorithms can readily be developed.

In particular, we specify the decision variables for the alternative formulation to be the number of vehicle trips scheduled in the pth period (denoted by N_p) and the number of passengers accommodated in this period (denoted by x_p).

The relationship between the variable N_p and the cumulative demand for service (D_p) that has been carried over to time period p is given by

$$D_p = \text{Max } \{h_p + D_{p-1} - \gamma N_p, 0\},$$

where, as before, h_p is the number of people entering the system at the beginning of the time period and γ is the capacity of each vehicle. With the demand quantities D_p thus identified, the number of passengers x_p accommodated in time period p must satisfy the inequality

$$x_p \leqslant \text{Min } \{h_p + D_{p-1}, \gamma N_p\}.$$

Thus, the objective (to maximize profits) may be written in the form:

$$\text{Maximize} \sum_{p=1}^{n} (c_p x_p - d_p N_p).$$

Here, once again, c_p denotes the passenger fare and d_p denotes the cost of the trip for vehicles departing in period p.

At first glance this latter representation of the problem may appear scarcely more tractable than the preceeding one, since the variables x_p depend nonlinearly on the variables N_p through the quantities D_p, which are themselves nonlinear functions of the N_p. However, as we will show, this representation can be substantially simplified by taking advantage of a few straightforward interconnections between variables and constraints. Although the problem still remains nonlinear, we will further show that it can be efficiently solved by a dynamic programming approach. Moreover, we will demonstrate that this approach can be adapted to solve problems with considerably more general interrelationships and more complex nonlinearities than those occurring in the prototype model.

3. A MATHEMATICAL PROGRAMMING FORMULATION

To develop a compact and useful mathematical programming formulation, we first make use of the previously stated assumption that the vehicle trip costs (d_p) are positive to strengthen the portion of the upper bound restriction on x_p that involves the variable N_p; i.e.,

$$x_p \leqslant \gamma N_p. .$$

In particular, in any optimal solution, the number of vehicle departures N_p that are scheduled in period p to carry x_p passengers can acceptably be specified to satisfy

$$N_p = \langle x_p / \gamma \rangle,$$

where the pointed brackets $\langle\ \rangle$ indicate the least integer greater than or equal to the quantity inside.

Thus, for example, if 150 passengers are given service in period p and the capacity of each vehicle is 100 passengers, then the number of departures in period p will be $N_p = \left\langle \frac{150}{100} \right\rangle = 2$. (Using more vehicles would obviously not make good economic sense, thus operating contrary to the goal of maximizing profit.) Consequently, by means of this relationship, the problem objective can be restated in the form

$$\text{Maximize} \sum_{p=1}^{n} (c_p x_p - d_p \langle x_p/\gamma \rangle).$$

This simplification in the objective function gives rise to a second simplification in the problem constraints. Specifically, the constraining relations of the model that are not accommodated by $N_p = \langle x_p/\gamma \rangle$ are precisely equivalent to $\sum_{p=1}^{i} x_p \leq \sum_{p=1}^{i} h_p$, $i = 1, \ldots, n$. These latter inequalities, which involve only the variables x_p and the constants h_p, succeed in eliminating the variables N_p from the problem entirely. They acceptably accommodate the restriction that the total *cumulative* number of passengers accommodated through each period i does not exceed the cumulative number that have arrived through that period. Thus, it is assured that the maximum number of passengers provided service in any given period is those available to be accommodated by that period.

The foregoing simplifications reduce the problem to an integer program with a nonlinear objective function and linear constraints. The linear constraints exhibit a special "stairstep" structure which is an instance of the "nested" structure studied in inventory theory by Ignall and Veinott [2] and in integer programming and network flows by Glover [3]. The success in devising highly efficient algorithms to accommodate problems that exhibit such constraints in other settings leads to the hope that such algorithms could also be devised in the present setting. We will show that this hope is not misplaced.

Indeed, by the indicated simplifications of the objective function and constraints, our prototype model can be solved by a direct adaptation of the methods of Bisbee [6] and Tingaud [7] for airline and railway scheduling models (reported in Uebe [5]). However, rather than specify such an adaptation, we shall instead consider a somewhat more general model structure which overcomes a variety of limitations in the assumptions underlying the prototype model. In this setting, our model more closely resembles the production planning model of Florian and Klein [9], except that we allow for nonconcavity in the objective function and also include nonlinearities in the constraints. Our main result specifies an efficient dynamic programming approach for this model that involves only single state variables. As will be proved, the exploitable feature of our problem is not the stairstep structure in a set of linear inequalities (which is implicitly or explicitly exploited in [2, 3, 6, 7, 9]), but rather the presence of this structure in the *arguments* of the problem functions (which may themselves be entirely nonlinear).

4. GENERALIZED PROBLEM STATEMENT AND SOLUTION METHOD

In this section we address ourselves to the problem of maximizing a separable nonlinear function

$$F_0(x) = \sum_{i=1}^{n} F_i\left(x_i, \quad \sum_{p=1}^{i} x_p\right),$$

subject to constraints

$$A_i \leq G_i\left(x_i, \sum_{p=1}^{i} x_p\right) \leq B_i \qquad \text{for all } i=1, \ldots, n \tag{4.1}$$

$$x_i \text{ integer,}$$

where A_i and B_i are vectors of constants and G_i is an arbitrary vector-valued function. The constraints (4.1) are assumed to contain or imply the constraints

$$a_i \leq \sum_{p=1}^{i} x_p \leq b_i$$

$$\text{for all } i=1, \ldots, n \tag{4.2}$$

$$v_i \leq x_i \leq u_i,$$

for some finite integer constants a_i, b_i, v_i, and u_i. (Note that the existence of finite a_i and b_i imply the existence of finite v_i and u_i, and vice versa. However, we include the full set of inequalities of (4.2) since they manifest a form that often arises in practical applications.)

The scheduling problem of the preceding section may be identified as a special instance of the current problem in which

$$F_i\left(x_i, \sum_{p=1}^{i} x_p\right) = c_i x_i - d_i \langle x_i/\gamma \rangle$$

and in which the inequalities of (4.1) reduce to those of (4.2) with $u_i = b_i = \sum_{p=1}^{i} h_i$ and $a_i = v_i = 0$. (The constants u_i and a_i are redundantly implied here by the constants b_i and v_i.)

5. DYNAMIC PROGRAMMING SOLUTION

To solve the problem of section 4 we introduce the related "parameterized" problem $P_k(z)$:

$$\text{Max} \sum_{i=1}^{k} F_i\left(x_i, \sum_{p=1}^{i} x_i\right)$$

subject to (4.1) and (4.2) for $i=1, \ldots, k$ and

$$\sum_{i=1}^{k} x_i = z.$$

Denoting the optimal objective function for $P_k(z)$ by $\alpha_k(z)$, we stipulate (by convention) that $\alpha_k(z) = -M$ if $P_k(z)$ has no feasible solution, where M is a "large number," such that $y - M = -M$ for all finite and infinite real numbers y.

Our main result may then be stated as follows.

THEOREM: Let $W_k(z) = \{\text{integer } w\colon v_k, z - b_{k-1} \leq w \leq u_k, z - a_{k-1} \text{ and } A_k \leq G_k(w, z) \leq B_k\}$.

Then, for $k=1,\ldots,n$, and $a_k \leq z \leq b_k$:

$$\alpha_k(z) = -M \text{ if } W_k(z) = \emptyset, \text{ and otherwise}$$

$$\alpha_k(z) = \operatorname*{Max}_{w \in W_k(x)} \{\alpha_{k-1}(z-w) + F_k(w,z)\},$$

where $a_0 = b_0 = 0$ and $F_0\ (w, z) = 0$ for all w, z.

PROOF: Clearly $\alpha_1(z)$ is correctly specified by the theorem for all z satisfying $a_1 \leq z \leq b_1$. Suppose then $\alpha_k(z)$ is correctly specified for all z satisfying $a_k \leq z \leq b_k$ and for all $k \leq r-1 < n$. Let z^* be a specific value of z constrained by $a_r \leq z^* \leq b_r$. The stipulation that $\alpha_r(z^*) = -M$ if $W_r(z^*) = \emptyset$ is obviously correct, and hence we suppose this condition does not apply. Since $w \epsilon W_r(z^*)$ only if $a_{r-1} \leq z^* - w \leq b_{r-1}$, it follows that $\alpha_{r-1}(z^*-w)$ is meaningfully defined for $w \epsilon W_r(z^*)$. Also, $\alpha_r(z^*)$ receives the value $-M$ from the dynamic programming recursion if and only if $\alpha_{r-1}(z^*-w) = -M$ for all $w \epsilon W_r(z^*)$. This, together with the previously noted circumstances under which $\alpha_r(z^*)$ can equal $-M$, precisely identifies the full range of conditions under which $P_r(z^*)$ can have no feasible solution. Thus, we henceforth restrict attention to those $w \epsilon W_r(z^*)$ for which $\alpha_{r-1}(z^*-w) > -M$ (if any exist). Since each such w can be made to provide a feasible solution to $P_r(z^*)$ by augmenting any feasible solution to $P_{r-1}(z^*-w)$ with the condition $x_r = w$, and since $\alpha_{r-1}(z^*-w)$ is the objective function value for a particular feasible solution to $P_{r-1}(z^*-w)$, it follows that $\alpha_r(z^*) \leq \Phi^*$, where Φ^* is the optimal objective function value for $P_r(z^*)$. But, in addition, if x^* denoted an optimal solution to $P_r(z^*)$, then the solution x' given by $x_i' = x_i^*$ if $i \neq r$ and $x_r' = 0$ provides a feasible solution to $P_{r-1}(z^* - x_r^*)$. Finally, by our previous observations, $x_r \epsilon W_r(z^*)$ and hence $\Phi' + F_r(x_r^*) = \Phi^* \leq \alpha_r(z^*)$, completing the proof.

6. CONCLUDING REMARKS

The dynamic programming recursion, stated in the previous theorem, can be applied by generating a "partial matrix" of the $\alpha_k(z)$ values (for the relevant range of the integer state variable z, given the stage k), together with the "optimal decisions" $w \epsilon W_k(z)$ that give rise to the entries of this matrix. Thus, the recursion successfully provides a method that effectively solves the prototype aspatial transportation model of [4], as well as a variety of problems with more general structures.

For example, the production planning model of Florian and Klein [9] is a restricted instance of this model in which the constraints of (4.1) reduce to the linear constraints of (4.2) with $a_i = v_i = 0$ and $b_i = \sum_{p=1}^{i} u_i$, and in which the function $F_i\left(x_i, \sum_{p=1}^{i} x_p\right)$ are given by $f_i(x_i) + g_i\left(\sum_{p=1}^{i} x_p - K_i\right)$, where f_i and g_i are concave and K_i is a constant.

In addition, as a simple extension of the model (presented in section 3), it is possible to allow the vehicle capacity to vary with the time period, and to permit the marginal demand (i.e., the incremental demand in each time period) to be a random variable. The latter change can easily be incorporated into the solution procedure—as long as the h_p's (in this case, the number of passengers entering the system in the pth time period) remain independent. A variety of other extensions compatible with the assumptions of the theorem will be indicated in a following paper.

Because of the demonstrated exploitability of the nested structure in the scheduling context, the results of the preceding sections suggest the attractiveness of finding reformulations for other schedul-

ing models (perhaps kindred to the reformulation developed here for the aspatial transportation model) that evoke such a structure from constraining relations in which it may not otherwise seem to be harbored.

REFERENCES

[1] Wagner, Harvey M., *Principles of Operations Research* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1969).

[2] Ignall, Edward and Arthur F. Veinott, Jr., "Optimality of Myopic Inventory Policies for Several Substitute Products," Management Science, *15*, 284–304 (1969).

[3] Glover, Fred, "Network Algorithms for Flows in Arborescences," *Proceedings of the Fifth International Conference on Operations Research*, Volume 10, No. 2, Venice (June 1969).

[4] Klastorin, T. D., "A Generalized Urban Transportation System," Department of Statistics and Operations Research, Graduate School of Business, The University of Texas, Austin (June 1970).

[5] Uebe, G., "Scheduling Regular Flights of an Airline," Mathematics Research Center, The University of Wisconsin, MRC Technical Summary Report #1132 (Feb. 1971).

[6] Bisbee, E. F., "Operations Analysis of System Specifications, Part I, Passenger Scheduling," MIT Report PB173635/R66–54.

[7] Tingaud, Jean-Piere, "On the Optimal Schedule of a One Line Railway System," Operations Research Center, University of California at Berkeley, Report ORC 68–6, March, 1968.

[8] Uebe, G., "Optimale Fahrplane," Springer, New York-Berlin-Heidelberg, 1970.

[9] Florian, Michael and Morton Klein, "Deterministic Production Planning with Concave Costs and Capacity Constraints," Départment d'Informatique, Université de Montréal (Dec. 1969) (revised Oct. 1970).

[10] Zangwill, W. I., "A Deterministic Multiperiod Production Scheduling Model with Backlogging," Management Science *13*, 105–119 (1966).

[11] Wagner, H. M., "A Postscript to Dynamic Problems of the Firm," Nav. Res. Log. Quart., 7, 7–12 (1960).

[12] Wagner, H. M. and T. M. Whitin, "Dynamic Version of the Economic Lot Size Model," Management Science *5*, 89–96 (1958).

FINDING EQUIVALENT TRANSPORTATION FORMULATIONS FOR CONSTRAINED TRANSPORTATION PROBLEMS*

F. Glover

University of Colorado, Boulder

Darwin Klingman

G. Terry Ross

University of Texas, Austin

ABSTRACT

This paper describes a procedure for determining if constrained transportation problems (i.e., transportation problems with additional linear constraints) can be transformed into equivalent pure transportation problems by a linear transformation involving the node constraints and the extra constraints. Our results extend procedures for problems in which the extra constraints consist of bounding certain partial sums of variables.

1.0. INTRODUCTION

The classical transportation problem is well known for its widespread applicability and for the facility with which it can be solved. Many procedures have been developed for reformulating disparate linear programming problems as transportation problems to take advantage of the computational efficiency inherent in the specialized transportation algorithm [2, 3, 8, 10]. We generalize the concepts of an earlier paper [6] to show how constrained transportation problems (i.e., transportation problems with additional linear constraints) can be transformed into pure transportation problems. Our procedure determines if an arbitrary extra linear constraint can be transformed into an equivalent bounded partial sum of variables involving a single node constraint. If this is possible the procedure gives the linear transformation that yields the equivalent constraint. This extends the work of Wagner [9], Manne [4, p. 382], and Charnes [1] who have shown how transportation problems with these bounded partial sums can be reformulated into pure transportation problems.

It is conjectured that our transformation requires computational effort on the same order as that required to find an initial basic solution by Vogel's Approximation Method. Since computational results indicate that specialized transportation codes can solve transportation problems at least 150 times faster than general purpose linear programming codes [5], our results make it possible to solve constrained transportation problems of the specified class with substantially greater efficiency than by a general purpose algorithm.

*This research was partly supported by a grant from the Farah Foundation and by ONR Contracts N00014–67–A–0126–0008 and N00014–67–A–0126–0009 with the Center for Cybernetic Studies, The University of Texas.

2.0. PROBLEM STATEMENT, MATHEMATICAL DEVELOPMENT, AND AN EXAMPLE

A transportation problem with one additional extra constraint (a singularly constrained transportation problem) can be stated mathematically as follows:

$$\text{Minimize} \sum_{i \epsilon M} \sum_{j \epsilon N} c_{ij} x_{ij},$$

subject to

$$\sum_{j \epsilon N} x_{ij} = a_i, \qquad i \epsilon M$$

$$\sum_{i \epsilon M} x_{ij} = b_j, \qquad j \epsilon N$$

$$\sum_{i \epsilon M} \sum_{j \epsilon N} p_{ij} x_{ij} = d$$

$$x_{ij} \geqslant 0, \qquad i \epsilon M, j \epsilon N,$$

where $M = \{1, 2, \ldots, m\}$, $N = \{1, 2, \ldots n\}$, $\sum_{i \epsilon M} a_i = \sum_{j \epsilon N} b_j$. In detached coefficient form the singularly constrained transportation problem appears as indicated in Table 1. The coefficients (p_{ij}) of the extra constraint appear below the familiar echelon-diagonal structure of ones of the transportation problem.

TABLE 1. *Detached Coefficient Form of Singularly Constrained Transportation Problems*

Coefficient of a linear combination	$x_{11}x_{12} \ldots x_{1n}x_{21}x_{22} \ldots x_{2n} \ldots x_{m1}x_{m2} \ldots x_{mn}$ $c_{11}c_{12} \ldots c_{1n}c_{21}c_{22} \ldots c_{2n} \ldots c_{m1}c_{m2} \ldots c_{mn}$		Variables Minimize
R_1	$1 \; 1 \ldots 1$	=	a_1
R_2	$1 \; 1 \ldots 1$	=	a_2
$\vdots$	$\vdots$	$\vdots$	
R_m	$1 \; 1 \ldots 1$	=	a_m
K_1	$1 \quad 1 \quad \ldots 1$	=	b_1
K_2	$1 \quad 1 \quad \ldots 1$	=	b_2
$\vdots$	$\ddots \quad \ddots \quad \ddots$	$\vdots$	
K_n	$1 \quad 1 \quad \ldots 1$	=	b_n
1	$p_{11}p_{12} \ldots p_{1n}p_{21}p_{22} \ldots p_{2n} \ldots p_{m1}p_{m2} \ldots p_{mn}$	=	d
Case A	$1 \; 1$	=	f
Case B	$1 \quad 1$	=	f
Case C	$1 \; 1$	=	f

Our goal is to specify a computationally simple procedure for identifying a linear combination of the ordinary transportation constraints, if one exists, which can be subtracted from the extra constraint to produce an inequality that is equivalent to establishing a bound for a partial sum of variables associated with a single origin or destination. From well known properties of the transportation matrix, if there exists a linear combination of the transportation constraints that has the desired form, then a linear combination can be found in which any particular transportation constraint receives a zero weight. Thus, we may arbitrarily delete a node constraint when seeking such a linear combination. Having done this certain variables can be viewed as having only one entry in the coefficient matrix for purposes of finding the desired linear combination. To exploit these facts in seeking a transformation of the singularly constrained transportation problem into an ordinary transportation problem, we arbitrarily omit the first origin constraint. Our principal observations are then contained in the following three cases.

CASE A. Assume that the extra constraint is equivalent to a partial sum of variables in a single origin constraint q other than origin 1 (i.e., the equivalent extra constraint is of the form $\sum_{j\epsilon S} x_{qj}=f$ for $S \subset N$ as illustrated in Table 1). Having set R_1 equal to 0, the unique values for all the destination constraint multipliers can be determined using the equations $R_1+K_j=p_{1j}$. This is possible by the assumption that the equivalent constraint involves variables in origin $q \neq 1$ (i.e., the coefficient on the variables x_{1j}, $j=1, \ldots, n$ in the equivalent constraint must equal 0). By the same reasoning a unique origin multiplier can be found for every other origin except origin q. In origin q not all of the equations $R_q+K_j=p_{qj}$ can simultaneously be satisfied by a single value for R_q. There will be two values, one that satisfies the equation for a subset $S \subset N$ and another that satisfies that equation for the destinations in the subset N–S. By setting R_q equal to the value $R_q+K_j=p_{qj}$ for $j\epsilon N$–S a linear combination of the standard node constraints has been found which when subtracted from the original extra constraint yields a restriction on the partial sum of variables in origin constraint q associated with the destinations in S. The equivalent constraint $\sum_{j\epsilon S} (p_{qj}-R_q-K_j)x_{qj}=d-\sum_{i\epsilon M} R_i a_i-\sum_{j\epsilon N} K_j b_j$ can be reduced to an equivalent partial sum by dividing through by $(p_{qj}-R_q-K_j)$ since this expression has the same value for all $j\epsilon S$.

CASE B. Assume that the extra constraint is equivalent to a partial sum of variables in a single destination q (i.e., the constraint is of the form $\sum_{i\epsilon T} x_{iq}=f$ for $T \subset M$ as illustrated in Table 1). The appropriate linear combination for this case can be found in the same manner as in Case A. After the origin constraint multiplier R_1 is set equal to zero and the destination constraint multipliers are determined using the equations $R_1+K_j=p_{1j}$, the remaining R_i values can be found. In every origin i that does not include variables in the equivalent partial sum (i.e., for $i\epsilon M - T$), there exists a unique value for R_i such that the constraints $R_i+K_j=p_{ij}$ will be satisfied simultaneously by a unique value for all $j\epsilon N$. For those origins $i\epsilon T$ that have a variable in common with the equivalent extra constraint, the equation $R_i+K_j=p_{ij}$ will be satisfied by a unique value for all $j \neq q$. Setting R_i, $i\epsilon T$ equal to this unique value will provide the scalar multipliers for a linear combination of the node constraints which when subtracted from the original extra constraint yield an equivalent partial sum of variables in the single destination constraint q associated with the origins in T. The equivalent constraint

$$\sum_{i\epsilon T}(p_{iq}-R_i-K_q)x_{iq}=d-\sum_{i\epsilon M}R_ia_i-\sum_{j\epsilon N}K_jb_j$$

can be reduced to an equivalent partial sum by dividing through by $(p_{iq}-R_i-K_q)$ since this expression has the same value for all $i\epsilon T$.

CASE C. Assume the extra constraint is equivalent to a partial sum of variables in origin constraint 1 (i.e. the constraint is $\sum_{j\epsilon S} x_{1j}=f$ for $S \subset N$ as illustrated in Table 1). Starting again with R_1 equal to zero, the K_j values for $j\epsilon N$ are immediately determined. Since the equivalent partial sum is assumed to be in origin 1, multipliers must be found to satisfy $R_i+K_j=p_{ij}$ for all i and j such that $i\neq 1$ in order for these variables not to appear in the equivalent constraint. However, given the current values for K_j unique values for R_i, $i\neq 1$ cannot be found, for if they could be then $R_i+K_j=p_{ij}$ would hold for *all* i and all j contradicting the assumption. Thus there must be two possible values for each R_i, $i\neq 1$. Set these R_i equal to the value that satisfies $R_i+K_j=p_{ij}$ for $j\epsilon N-S$. Thus, $R_i+K_j\neq p_{ij}$ for $i\neq 1$ and for $j\epsilon S$. Note, however, new values for K_j can be found that satisfy $R_i+K_j=p_{ij}$ for $i\neq 1$ and $j\epsilon S$. Thus by changing the appropriate K_j values after having determined values for all R_i, one can find the linear combination of the node constraints which when subtracted from the extra constraint leaves the equivalent restriction $\sum_{j\epsilon S}(p_{1j}-K_j)x_{1j}=d-\sum_{i\epsilon M}R_iA_i-\sum_{j\epsilon N}K_jb_j$. Since $(p_{1j}-K_j)$ is the same for all $j\epsilon S$ then dividing through by this quantity will yield the equivalent partial sum.

Based on the reasoning just presented we can describe a general procedure to effect the transformation as required in the three cases. It is convenient to use a transportation tableau with the coefficients of the extra constraint placed in the cell corresponding to the appropriate variable. This format is illustrated in Table 2. Values for the R_i and K_j multipliers are shown around the rim of the tableau and can be determined in a manner similar to that used to find values for dual evaluators for a basic solution to a transportation problem.

TABLE 2. *Tableau Representation of an Extra Constraint*

	K_1	K_2	$\cdots$	K_n	
R_1	p_{11}	p_{12}	$\cdots$	p_{1n}	a_1
R_2	p_{21}	p_{22}	$\cdots$	p_{2n}	a_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
R_m	p_{m1}	p_{m2}	$\cdots$	p_{mn}	a_m
	b_1	b_2	$\cdots$	b_n	

Our procedure for finding the appropriate linear combination of the constraints is:

STEP 1. Set R_1 equal to zero.

STEP 2. For all $j\epsilon N$ set K_j equal to p_{1j}.

STEP 3. Try to determine a unique value for R_2 using the equations $R_2+K_j=p_{2j}$ for all $j\epsilon N$.

a) If a unique value for R_2 can be found set R_2 equal to this value and proceed to step 4.

b) If the equations are not satisfied by a unique value of R_2 but all equations except q can be satisfied by a single value of R_2, set R_2 so that the $R_2+K_q \neq p_{2q}$ and mark cell $(2, q)$ with a "star". However, if this is not the case, and R_2 must assume two distinct values to satisfy the equations for all j, set R_2 equal to either of two values arbitrarily and "star" the cells for which $R_2+K_j \neq p_{2j}$. Proceed to step 4.

c) If more than two values of R_2 are required to satisfy all the equations, stop. The constraint is not equivalent to a partial sum of variables in a single node constraint.

STEP 4. Continue determining values for the remaining R_i as in step 2 except when R_i must assume two distinct values in order for the equations $R_i+K_j=p_{ij}$ to be satisfied for all j. In this case set R_i so that the starred cells in this row lie in the same columns as those starred in earlier rows. If this cannot be done, stop. The constraint is not equivalent to a partial sum of variables in a single node constraint. Also for any column r with starred cells check to see that $p_{ir}-R_i=p_{kr}-R_k$ for $i \neq k$ where i and k are rows containing the starred cells. If this is not the case for all $i \neq k$, stop. Again the constraint is not equivalent to a partial sum of variables in a single node constraint.

STEP 5. After all R_i have been determined, four cases are possible.

i) the starred cells occur only in a single row.

ii) the starred cells occur in all cells in a subset of the columns except for the cells in row 1.

iii) the starred cells occur only in a single column.

iv) the starred cells occur in some cells in a subset of the columns, but not in row 1 and not in some other row.

In case i the starred cells indicate the variables included in the equivalent partial sum of variables. For these starred cells the coefficient on these variables in an equivalent constraint is $p_{ij}-R_i-K_j$. After forming the linear combination from the extra constraint only those terms will remain. Since $p_{ij}-R_i-K_j$ will be the same for all starred cells, the equivalent partial sum can be obtained by dividing through by the coefficient on the variables.

In case ii the K_j values for the columns containing starred cells can be changed so that the equation $R_i+K_j=p_{ij}$ holds for the starred cells. The effect is to "erase" the stars from these cells and place them in the cells in row 1 in these columns. This is now the same as case i.

In case iii the starred cells indicate the variables in a single destination constraint that comprise an equivalent constraint. Since $p_{ij}-R_i-K_j$ will be the same for all starred cells the equivalent partial sum can be obtained by dividing through by the coefficient of the variables.

In case iv the constraint is not equivalent to a partial sum of variables in a single origin or destination constraint.

For cases i, ii, and iii we have found the coefficients of the variables in an equivalent constraint. The new right-hand side value can be found by subtracting $\sum_{i \epsilon M} R_i a_i + \sum_{j \epsilon N} K_j b_j$ from the original right-hand side.

We can summarize our results to this point in the following theorem.

THEOREM: If an extra constraint is equivalent by a linear transformation to a partial sum of variables in a single node constraint then the stated procedure finds the equivalent partial sum.

PROOF: The hypothesis allows us to assert that the equivalence can be determined by a linear combination of the node constraints of the transportation problem. The remainder of the proof is contained in cases A, B, and C above.

A four-origin, five-destination constrained transportation problem is shown in Table 3 with the coefficients of the extra constraint indicated in the cells corresponding to the appropriate variables. The associated supply and demand values are shown along the rim of the table. Assume that the extra constraint is an "equality" constraint with a right-hand side value of 84.

TABLE 3

						Supply
	0	1	0	4	−5	10
	6	7	4	8	−1	15
	5	6	5	9	0	10
	−1	0	−1	3	−6	10
Demand	8	7	9	6	15	

By applying the procedure described above, the following R_i and K_j multipliers are obtained.
In step 1 R_1 is set equal to zero.
In step 2 the K_j multipliers are set equal to the following values

$$K_1=0, \qquad K_2=1, \qquad K_3=0, \qquad K_4=4, \qquad K_5=-5.$$

In step 3 a unique value for R_2 cannot be found to satisfy $R_2+K_j=p_{2j}$ for all j. The value 6 satisfies that equation for p_{21} and p_{22} and the value 4 satisfies that equation for p_{23}, p_{24}, and p_{25}. Arbitrarily set R_2 equal to 6 and star the cells p_{23}, p_{24}, and p_{25}.

As required by step 4 R_3 is set equal to 5 because this value satisfies the equation $R_3+K_j=P_{3j}$ for all j. Similarly a value of $R_4=-1$ satisfies $R_4+K_j=P_{4j}$ for all j. The results of the procedure to this point are shown in Table 4.

TABLE 4

	$K_1=0$	$K_2=1$	$K_3=0$	$K_4=4$	$K_5=-5$	Supply
$R_1=0$	0	1	0	4	−5	10
$R_2=6$	6	7	4*	8*	−1*	15
$R_3=5$	5	6	5	9	0	10
$R_4=-1$	−1	0	−1	3	−6	10
Demand	8	7	9	6	15	

To complete the transformation, next compute the values of $(p_{ij}-R_i-K_j)$ for the starred cells, check that they are equal, and obtain the coefficients for the variables in an equivalent constraint. In particular the left-hand side of the equivalent constraint is $-2x_{23}+-2x_{24}+-2x_{25}$. The new right-hand side value is found by the formula $(+84)-\sum_{i=1}^{4} R_i a_i-\sum_{j=1}^{5} K_j b_j=84-(130)-(-44)=-2$. Dividing through by -2, we obtain the equivalent partial sum $x_{23}+x_{24}+x_{25}=1$. Thus, by using the procedure described above we have found a partial sum of variables in a single row equivalent to the original

extra constraint. The original transportation problem can be enlarged by one source and one destination in the manner suggested by Wagner [9], and an optimal solution to the original problem can be found using the transportation algorithm.

3.0. EXTENSIONS

In the development of the procedure for transforming extra constraints and in the example the original extra constraint was assumed to be an "equality" type. It should be clear that the same transformation can be made for both "less than or equal" and "greater than or equal" constraints. If there are several extra constraints then the procedure can be applied to each one separately to obtain an equivalent partial sum for each extra constraint. Wagner [9] has shown that if these partial sums involve disjoint sets of variables and if the sets are nested in the same node constraints then the problem can be transformed into an enlarged transportation problem.

4.0. APPLICATIONS

Many models have the structure of a transportation problem with additional restrictions. The extra constraints may represent secondary objectives or restrictions that are not reflected in the standard node constraints. To illustrate a typical example of this class of problems, consider the transportation model where warehouses supply markets and the objective is the standard one of finding a shipping pattern that will minimize the total shipping cost. Suppose additionally that the products shipped are of a fragile nature and if they are sent via particular routes each item must be specially packaged to prevent losses in shipping. Table 5 shows the packaging time in minutes per unit required to prepare a unit for shipment on the various routes. Suppose that we wish to limit the average packaging time per unit to at most 2½ minutes per unit (i.e., for the 60 units that must be shipped, the total packaging time must not exceed 150 minutes). Using our procedure a linear combination of the node constraints of the transportation problem can be found which when subtracted from the original extra constraint yields an equivalent partial sum of variables $x_{22}+x_{24} \leq 5$. The node constraint multipliers for the linear combination are indicated in Table 5. Thus the 2½-minute average packaging time restriction can only be satisfied if the total number of units shipped along routes (2, 2) and (2, 4) is less than or equal to 5. The transportation problem can be transformed to include this restriction directly, and thus the original constrained transportation problem can be solved as a transportation problem with one additional origin and one additional destination.

TABLE 5

	$K_1=0$	$K_2=1$	$K_3=0$	$K_4=2$	$K_5=1$	Supply	
$R_1=0$	0	1	0	2	1	10	
$R_2=1$	1	3*	1	4*	2	20	
$R_3=3$	3	4	3	5	4	20	
$R_4=2$	2	3	2	4	3	10	
Demand	15	10	5	5	25	60	Total

BIBLIOGRAPHY

[1] Charnes, A. Lecture Notes for Extremal Methods at Carnegie-Mellon, Purdue, Northwestern University and The University of Texas at Austin.

[2] Charnes A., F. Glover and D. Klingman, "The Lower Bounded and Partial Upper Bounded Distribution Model," Nav. Res. Log. Quart., *18*, 277–281. (1971).

[3] Charnes, A. and D. Klingman, "The Distribution Problem with Upper and Lower Bounds on the Node Requirements," Management Science, *16*, 638–642 (1970).

[4] Dantzig, G., *Linear Programming and Extensions*, Princeton University Press, 1963.

[5] Glover, F., D. Karney, D. Klingman, and A. Napier, "A Computation Study on Start Procedures, Basis Change Criteria, and Solution Algorithms for Transportation Problems," To appear in Management Science.

[6] Glover, F., D. Klingman, and G. T. Ross, "Equivalence of Boolean Constrained Transportation Problems to Transportation Problems," Research Report CS 106, Center For Cybernetic Studies, The University of Texas, November, 1972.

[7] Klingman, D. and R. Russell, "On Solving Constrained Transportation Problems," Research Report CS 96, Center For Cybernetic Studies, The University of Texas, August 1972.

[8] Orden, A., "The Transhipment Problem," Management Science, *2*, 276–285 (1956).

[9] Wagner, H., "On a Class of Capacitated Transportation Problems," Management Science, *5*, 304–318 (1959).

[10] Wagner, H., "The Lower Bounded and Partial Upper Bounded Distribution Model," Nav. Res. Log. Quart. *20*, 265–268, (1973).

COMPUTERIZED SCHEDULING OF SEAGOING TANKERS*

M. D. McKay† and H. O. Hartley

Texas A&M University

ABSTRACT

Computerized Scheduling of Seagoing Tankers

The tanker scheduling problem considered in this paper is that of the Defense Fuel Supply Center (DFSC) and the Military Sealift Command (MSC) in the worldwide distribution of bulk petroleum products. Routes and cargoes which meet delivery schedule dates for a multiplicity of product requirements at minimum cost are to be determined for a fleet of tankers. A general mathematical programming model is presented, and then a mixed integer model is developed which attempts to reflect the true scheduling task of DFSC and MSC as closely as possible. The problem is kept to within a workable size by the systematic construction of a set of tanker routes which does not contain many possible routes that can be judged unacceptable from practical considerations alone.

1. INTRODUCTION

This paper represents a condensed summary of the mathematical aspects of a project which was carried out with the active cooperation of

(i) Defense Supply Agency Headquarters (DSAH),
(ii) Defense Fuel Supply Center (DFSC),
(iii) Military Sealift Command (MSC), and
(iv) Office of Naval Research (ONR).

It must not be regarded as a complete statement of the logistic ramifications of the operational problem of the scheduling of tankers arising at (i), (ii), (iii), but rather as a somewhat idealized version of the "real life" situation. We should state, however, that the computer program we have developed, which is briefly sketched in sections 5 and 6, has been made available for operational use to (ii).

Briefly speaking, the general problem of "tanker scheduling" is as follows: For each of a number of petroleum products (p) we are given assets Z_{jp} (in barrels) available at ports (j). Likewise we are given barrels X_{jp} of products (p) required at port (j) for a number of (destination) ports scheduled to arrive within specified time periods, and finally a fleet of tankers (τ) with given volume capacities C_τ (barrels) and weight limits W_τ (tons) and speeds v_τ (nautical miles per day). It is then required to solve the constrained transportation problem of shipping some or all of the available products by some or all of the tankers to the required destinations to arrive within the specified time limits. The problem involves both the selection of the tankers of adequate capacity and tonnage as well as the determination of their routes commencing at given "ports of availability" at given "dates of availability." The

*An earlier version of this paper was presented by H. O. Hartley at the Conference on Transportation Planning and Scheduling sponsored by ONR in George Washington University, March 17–19, 1971.

†Now at the Los Alamos Scientific Laboratory, University of California, Los Alamos, N.M. 87544.

objective of the tanker scheduling is then to minimize the total of the transportation costs (involving differential costs of running tanker τ for a day) plus the "purchasing cost" of the product (involving differential prices of product p at source ports j).

Before turning to the mathematical formulations of the problem in sections 4 and 5 we are surveying the limited literature dealing with special cases.

2. RELATION TO PUBLISHED WORK

We confine ourselves here to some references to papers dealing with the application of mathematical programming to the problem of tanker scheduling. The earliest paper is that by Dantzig and Fulkerson [5] who reduce the problem of minimizing the number of identical tankers satisfying a number of requirements (which are integer multiples of the tanker capacity) each from one single port, to a Hitchcock Transportation problem. The opposite problem in which the number of tankers available is insufficient to meet the requirements is considered by Bellmore, et al. [3]. These authors maximize the "utility" of the deliveries made by the tankers by using the Fulkerson "out of kilter algorithm," see Ford and Fulkerson [7]. Again the requirements are multiples of the capacity of identical tankers, no partial deliveries are allowed and tankers must after delivery go to a port for a next delivery or cease operating. Multiple products are allowed. In a subsequent paper by Bellmore, et al. [4] extensions are introduced, but the optimization of the utility may not be integral. The problem is formulated in terms of arc flow variables in a network and is solved by the Dantzig-Wolfe [6] decomposition algorithm. An enumerative branch and bound algorithm is proposed to reach integer solutions. Applegren (1969) considered a more realistic problem, but he still assumed fixed cargoes that were the same size as the ships, and single port discharges.

It will be seen that the problem here formulated is more general than that considered by Bellmore et al. [4]. The most important generalizations are:

(1) Multiple deliveries of multiple products at several destinations and lifting of products at more than one port to form cargoes are allowed for.

(2) Deliveries must occur within the scheduled delivery times.

(3) Our problem is solved for individual tankers starting at individual locations at different times.

(4) Differential port and product purchasing costs are involved in the utilities.

(5) Cargoes are not preassigned, but are determined as part of the solution.

It should also be noted that our problem is not tanker constrained: The available tanker fleet is assumed to be adequate* to satisfy the requirements, and the cost of achieving this is minimized.

3. COMPLETE NOTATION FOR PROBLEM FORMULATION

The subscript $j=1, 2, \ldots, J$ will be used to index ports. The first J_S ports are source ports and the remaining ports are destination ports. When it is not necessary to distinguish between source ports and destination ports, the term "location" may be used.

The following subscript notations will be used:

(1) j = location number

p = product number

---

*Strategies of "spot chartering" of commercial tankers cannot be discussed here.

τ = tanker number
l = "stop" number
n = occurrence number for product at location.

Since each product p could be required or available more than once at a location j, the subscript combination jp should be jpn in the development to follow. However, to make the description more understandable, the n subscript will be omitted.

In connection with the tankers, the following is assumed available:

(2) C_τ = capacity of tanker τ in barrels
W_τ = capacity of tanker τ in tons
v_τ = speed of tanker τ in nautical miles per day
b_τ = cost per day of running tanker τ
$j(\tau)$, $T(\tau)$ = location number at which tanker τ is expected to be available on calendar date $T(\tau)$.

For products, requirements, and assets the necessary information is

(3) X_{jp} = barrels of product p required at location j between calendar dates s_{jp} and S_{jp}
Z_{jp} = barrels of product p available at location j
a_{jp} = cost per barrel associated with asset Z_{jp}.

4. INTEGER PROGRAMMING FORMULATION

The solution of the integer programming problem will involve determining the values of two sets of variables. The tanker routes are specified by the 0–1 integer variables

$$u_{j\tau l}=\begin{cases}1 \text{ if tanker } \tau \text{ makes its } l\text{th stop at location } j\\ 0 \text{ if not,} \qquad j,\tau,l=1,2,\ldots.\end{cases} \tag{4}$$

The initial known location $j(\tau)$ will be denoted by $l=0$. Hence,

$$u_{j(\tau)\tau 0}=1 \qquad \text{and } u_{j\tau 0}=0 \qquad \text{for } j\neq j(\tau). \tag{5}$$

Moreover, certain port restrictions require that certain prespecified $u_{j\tau l}=0$ for all l.

The number of stops which tanker τ will make is to be determined by imposing the linear inequalities

$$1\geqslant\sum_j u_{j\tau 1}\geqslant\sum_j u_{j\tau 2}\geqslant\ldots\geqslant\sum_j u_{j\tau l}\geqslant\ldots\sum_j u_{j\tau L^*}\geqslant 0, \qquad \tau=1,2,\ldots \tag{6}$$

where L^* is a suitably chosen number (realistically of the order 5 to 7) representing the maximum number of stops per tanker and depending on the planning horizon.

In terms of the $u_{j\tau l}$ it is possible to express the expected arrival time $T_{\tau l}$ of tanker τ at its lth stop.

This is given by

$$T_{\tau l}=\frac{1}{v_\tau}\sum_{t=0}^{l-1}\sum_{jj'} u_{j\tau t}u_{j'\tau t+1}d_{jj'}+(l-1)\cdot\delta+T(\tau), \qquad \tau, l,=1, 2, \ldots \tag{7}$$

where,

$$d_{jj'}=\text{distance in nautical miles from location } j \text{ to location } j' \tag{8}$$

and

$$\delta=\text{average delay at each stop for loading and unloading.}$$

The tankers' intake and discharge, and hence the cargoes, are determined by the variables

$$z_{\tau pl}=\text{barrels of product } p \text{ loaded into tanker } \tau \text{ at the location of its } l\text{th stop}$$

$$x_{\tau pjl}=\text{barrels of product } p \text{ discharged by tanker } \tau \text{ at stop } l \text{ at location } j. \tag{9}$$

These variables must be constrained on tanker capacity. Also, it is often desirable to ensure that the tanker is filled to an adequate proportion q of its capacity after leaving each source port in its route. Thus, it is required that

$$q\Big(\sum_{j\leq J_S} u_{j\tau l}\Big)C_t\leq\sum_p\sum_{t=0}^{l} z_{\tau pt}-\sum_{pj}\sum_{t=0}^{l} x_{\tau pjt}u_{j\tau t}\leq C_\tau, \qquad \tau, l=1, 2, \ldots. \tag{10}$$

The restrictions on tonnage can be imposed by replacing C_τ, the volume capacity, by W_t the weight capacity, and by multiplying $z_{\tau pt}$ and $x_{\tau pjt}$ by the density of product p in the inequalities above.

It must also be required that for each product p the cumulative balance on board must be non-negative at each stop l. This means that

$$\sum_{t=0}^{l} z_{\tau pt}-\sum_j\sum_{t=0}^{l} x_{\tau pjt}u_{j\tau t}\geq 0, \qquad l, \tau, p=1, 2, \ldots. \tag{11}$$

Now, in order to ensure that the requirements are met and delivered during the prescribed times, the constraints

$$\sum_{\tau l} x_{\tau pjl}u_{j\tau l}=X_{jp}, \qquad j, p=1, 2, \ldots,$$

and

$$(T_{\tau l}-s_{jp})\, x_{\tau pjl}\geq 0 \tag{12}$$

$$(S_{jp}-T_{\tau l})x_{\tau pjl}\geq 0, \qquad \tau, p, j, l=1, 2, \ldots$$

are used. Finally, so that availabilities are not exceeded,

$$\sum_{\tau l} z_{\tau pl} u_{j\tau l} \leq Z_{jp}, \qquad j, p = 1, 2, \ldots \tag{13}$$

must be satisfied.

The objective function is a composite cost of running the tankers for the duration of their trips plus the cost of purchasing the products p at the port j. Therefore it is desired to minimize

$$F(u, z) = \sum_{\tau} \sum_{l=1}^{L^*} \left(\sum_{j} u_{j\tau l} \right) (T_{\tau l} - T_{\tau l-1})\, b_{\tau} + \sum_{\tau plj} a_{jp} z_{\tau pl} u_{j\tau l}. \tag{14}$$

It will be noticed that the problem formulated above is linear in the variables $z_{\tau pl}$ and $x_{\tau pjl}$ for *given* integer variables $u_{j\tau l}$. This suggests that the *conditional* minima of $F(u, z)$ given by (14) for given tanker routings $u_{j\tau l}$ can be determined by linear programming. However, the computation of all possible conditional minima is not feasible since theoretically the number of all possible routings for T tankers is $J^{T \times L^*}$ a clearly astronomical figure! On the other hand, a study of the requirement schedules s_{jp}, S_{jp} defined in Equation (3) revealed that in all practical situations encountered for a given tanker τ it is easy, a priori, to determine the small fraction of J^{L^*} possible routes which are feasible by ascertaining that all their L^* stops at destinations j meet the products p scheduled delivery times s_{jp}, S_{jp}.

5. RESTRUCTURED FORMULATION

The model of the last section is structured so that a tanker, at least initially, may make any sequence of stops. The number of potential routes for each tanker is of the order of J^{L^*} (around 10^7 for a typical DFSC task). Almost all of these routes are unacceptable for a variety of reasons. The two occurring most frequently are early or late arrivals at discharge ports, and stops at discharge ports when the product required could not possibly be on board the ship. Although the timeliness of the arrivals is controlled by constraints (12) on the problem, the difficulty in finding a solution is increased because of the quadratic occurrences of the variables in those constraints.

It was noted in the last section that if a single route is chosen for each tanker, the problem of determining the associated cargoes becomes a linear programming problem. The procedure to be proposed here is that instead of allowing any of the possible routes to be assigned to a tanker, a smaller set of "acceptable" routes be constructed, and that it be required in the solution strategy that only routes in that set be considered.

Let R denote a set of routes for the tanker fleet. It will be shown in a later section how R can be constructed. For now, it is adequate to assume that the routes in R have been generated by examining requirements, assets, and fleet configuration, in conjunction with tanker speeds and capacities, distances, port restrictions, etc. Hence, for each route in R it is known what requirements can be (partially) satisfied and what assets are to be used to (partially) satisfy them. Also, all deliveries will be timely; for if they *were not* timely, the route is not in R. The scheduling model to be presented now assumes the existence of the set R.

Restatement of the Problem and Notation

The routes for tanker τ are labeled by the 0–1 variables $q_{\tau r}$, where

(15) $$q_{\tau r} = \begin{cases} 1 \text{ if the } r\text{th route of tanker } \tau \text{ is selected} \\ 0 \text{ if not}, \qquad \tau, r = 1, 2, \ldots . \end{cases}$$

The product assets will be indexed by $\alpha = 1, 2, \ldots$, and the product requirements by $\rho = 1, 2, \ldots$. The specification of a value for α carries with it an implicit specification of a port where the product is available, the contract, amount, etc. Likewise, each value of ρ implies a port, product, amount, and delivery dates. Use will be made of this implicit specification, making subscripts pointing to particular ports and products unnecessary. The variables to be determined in addition to the $q_{\tau r}$'s are

(16) $x_{\alpha\rho\tau r}$ = amount of asset α sent to satisfy requirement ρ on the rth route of tanker τ.
p_τ = fraction of the volume capacity of tanker τ used.

An examination of the routes in R will show that many of the $x_{\alpha\rho\tau r}$ must be zero, because for the given route r no fraction of asset α can be shipped to satisfy requirement ρ within the delivery dates. A monitoring algorithm based on each route will eliminate the impossible $x_{\alpha\rho\tau r}$ from the problem.

Also to be used are

(17)
X_ρ = amount of requirement ρ
Z_α = amount of asset α
a_α = cost per barrel associated with asset α
C_τ = volume capacity of tanker τ
W_τ = weight capacity of tanker τ
b_τ = cost per day of operating tanker τ
$T_{\tau r}$ = length in days of the rth route of tanker τ.

Again, $T_{\tau r}$ is obtained upon examination of R.

In the formulation of the general problem in the preceding section, Equalities (12) ensured that all requirements are, in fact, satisfied if a feasible solution exists. Rather than have a model whose solution may point up infeasibilities, the one described below will "do its best" to satisfy the requirements, and certainly satisfy them all if a feasible solution exists. This is achieved by setting upper bounds, namely the amount required, for the amounts delivered. A "reward," or utility, given for meeting the requirements will be added to the objective function. In light of this, the parameters

(18) M_ρ = reward per unit for satisfying requirement ρ

H_τ = reward per unit for filling tanker τ

will be used.

Formulation

The objective function to be maximized, the rewards for satisfying requirements for filling tankers

less the cost for doing so, is given by

$$f(x, p, q) = \sum_{\alpha\rho\tau r} M_\rho x_{\alpha\rho\tau r} - \sum_{\alpha\rho\tau r} a_\alpha x_{\alpha\rho\tau r} + \sum_{\tau} H_\tau p_\tau - \sum_{\tau r} b_\tau T_{\tau r} q_{\tau r}. \tag{19}$$

The constraints on the solution are as follows. To ensure that requirements are not *over* satisfied, and that availabilities are not overrun, it is required that

$$\sum_{\alpha\tau r} x_{\alpha\rho\tau r} \leq X_\rho, \qquad \rho = 1, 2, \ldots \tag{20}$$

and

$$\sum_{\rho\tau r} x_{\alpha\rho\tau r} \leq Z_\alpha, \qquad \alpha = 1, 2, \ldots.$$

Since each tanker τ can be sent on at most one route, it is further required that

$$\sum_{r} q_{\tau r} \leq 1, \qquad \tau = 1, 2, \ldots \tag{21}$$

and

$$q_{\tau r} \epsilon \{0, 1\}, \qquad \tau, r = 1, 2, \ldots.$$

These two constraints limit at most one $q_{\tau r}$ to 1 for each tanker τ.

The maximum capacity that can be allocated to route r of tanker τ is $q_{\tau r} C_\tau$. Hence, to limit the amounts loaded, it must be that

$$\sum_{\alpha\rho} x_{\alpha\rho\tau r} \leq q_{\tau r} C_\tau, \qquad \tau, r = 1, 2, \ldots. \tag{22}$$

The variable p_τ becomes the fraction capacity utilization of tanker τ by virtue of the objective function and the constraint

$$\sum_{\alpha\rho r} x_{\alpha\rho\tau r} \geq p_\tau C_\tau, \qquad \tau = 1, 2, \ldots \tag{23}$$

because from (21), (22), and (23) the $x_{\alpha\rho\tau r}$ can be nonzero for at most one route r for each tanker, and

$$\sum_{r} q_{\tau r} \geq p_\tau, \qquad \text{for each } \tau. \tag{24}$$

In order to ensure that draft restrictions of the ports are met by incoming tankers, introduced into the model are constraints of the form

$$D_\tau - Y_{\tau r l}\rho_\tau \leq d_{\tau r l}, \tag{25}$$

where

D_τ = Dead weight draft in feet of tanker τ,

$d_{\tau r l}$ = Draft of the lth port of route r for tanker τ,

ρ_τ = Conversion factor from barrels to feet for tanker τ, by using an average product density,

and

$Y_{\tau r l}$ represents a sum over certain $x_{\alpha\rho\tau r}$ to yield the cargo discharged after leaving the last source port before the discharge port in question.

The inequalities above can be rewritten as

$$\sum_r Y_{\tau r l} \geq \sum_r q_{\tau r}(D_\tau - d_{\tau r l})/\rho_\tau, \qquad \tau, l = 1, 2, \ldots. \tag{26}$$

Since, for each τ at most one $q_{\tau r}$ is one (and the rest are zero), and the $Y_{\tau r l}$, really a sum of $x_{\alpha\rho\tau r}$'s, corresponding to that $q_{\tau r}$ is the only one that is nonzero by virtue of (22). The advantage of (24) over (25) is that the number of inequalities is independent of the number of routes for the tanker.

Finally, the nonnegativity constraints,

$$x_{\alpha\rho\tau r} \geq 0, \qquad \alpha, \rho, \tau, r = 1, 2, \ldots \tag{27}$$

are required.

Construction of the Set R of Acceptable Routes

The details of the selection rules used to generate routes which meet DFSC and MSC specifications will not be presented here, but rather a brief outline of the procedure followed.

For each source port a table of destination ports which *could* be serviced from that port is constructed by examining the assets available at the source port and the requirements of the destination port. The maximum amounts (in barrels) that could be shipped from the source to the destinations are also recorded. These quantities, called Maximum Serviceable Amounts (MSA), are used later to obtain lower bounds on maximum tanker capacity utilization and tanker draft. It may be desired to prohibit certain source-destination combinations, for example to limit imports into the United States, during the construction of the tables.

Then taking tankers individually and using the tables, MSA, tanker speed and draft, and port draft, sequences of port calls which could be feasible routes are constructed. For the DFSC/MSC problem, a route was defined as calls at one or two loading ports followed by calls at up to three destination ports.

6. SOLUTION TO THE INTEGER PROGRAMMING PROBLEM

For a typical DFSC task the integer programming problem of section 5 contained 700–900 integer variables, 2,500 continuous variables, and 1,000 constraints. It was felt that to try to solve the problem completely would not be (economically) feasible, hence an approximate solution technique was used.

The problem was set up without the integerality constraints on the 0–1 variables $q_{\tau r}$ in constraints (21). The dropping of the integer constraint is equivalent to "splitting" each tanker into "fractional tankers" with a total capacity equal to that of the original tanker. In the resulting linear programming problem, the variables $q_{\tau r}$ become the fraction of the capacity C_τ of tanker τ allocated to its rth route.

In order to produce a solution to the *LP* problem where the capacity for each tanker is allocated to at most one route, an iterative solution scheme with "selective" rounding was employed. To facilitate this technique, the inequality

$$\sum_{\tau r} d_{\tau r} q_{\tau r} \geq N \tag{28}$$

with $d_{\tau r}$ and N initially set to zero, was introduced into the model. This *single* inequality is used in conjunction with

$$\sum_{r} q_{\tau r} \leq 1, \qquad \tau = 1, 2, \ldots \tag{21}$$

to force at most one $q_{\tau r}$ to 1 for each tanker τ. One way to maintain a $q_{\tau r}$ at 1 is to use parametric programming to bring the $q_{\tau r}$ to 1, and then to introduce a lower bound constraint into the problem to keep it there. However, it was found that the use of the procedure to be described presently was much more efficient in terms of computer run time.

An iteration consisted of the following steps:

(1) Solve the existing *LP* problem.

(2) Examine the set of $q_{\tau r}$ in the optimal basis, and select the largest $q_{\tau r}C_\tau$, the total capacity utilized by tanker τ on route r, for which $d_{\tau r} = 0$. (Recall that initially all $d_{\tau r} = 0$.)Let $q_{\tau^* r^*} C_{\tau^*}$ denote the selected capacity. If $q_{\tau^* r^*} = 0$, the selection process terminates. That is no more tankers will be routed. If $q_{\tau^* r^*} > 0$, modify the existing *LP* by setting $d_{\tau^* r^*} = 1$ and increasing N by 1. Then begin again at (1).

When the iteration loop terminates, i.e., for each $q_{\tau r}$ in the optimal basis either $q_{\tau r} = 0$ or $d_{\tau r} \neq 0$, all "final" routes are examined. If the tanker is 80 percent full after its first lift port, or if an increase in its cargo by 10 percent could bring it up to 80 percent filled capacity, the route is accepted as part of the final scheduling.

The transition from Step (2) of one iteration to Step (1) of the next involves the modification of 1 element in the *LP* coefficient matrix and 1 element in the right hand side. To gain advantage of the previous optimal solution when solving the next *LP*, a modification of the existing basis is made in order to make it a feasible basis for the modified *LP* problem. Since the only infeasibilities that can occur arise in the Inequalities (21), only minor modification need be made.

REFERENCES

[1] Applegren, L. H., "A Column Generation Algorithm for a Ship Scheduling Problem," Transportation Science *3*, 53–68 (1969).

[2] Applegren, L. H., "Integer Programming Methods for a Vessel Scheduling Problem," Transportation Science *5*, 64–78 (1971).

[3] Bellmore, M., G. Bennington and S. Lubore, "A Maximum Utility Solution to a Vehicle Constrained Tanker Scheduling Problem," Nav. Res. Log. Quart. *15*, 404–411 (1968).

[4] Bellmore, M., G. Bennington and S. Lubore, "Further Extensions of the Tanker Scheduling Problem," Presented at the 36th National Operations Research Society Meeting in Miami, Florida, November 1969.

[5] Dantzig, G. B. and D. R. Fulkerson, "Minimizing the Number of Tankers to Meet a Schedule," Nav. Res. Log. Quart. *1*, 217–222 (1954).

[6] Dantzig, G. B. and P. Wolfe, "Decomposition Principle for Linear Programs," Operations Research *8*, 101–111 (1960).

[7] Ford, L. R. and D. R. Fulkerson, "*Flows in Networks*" (Princeton University Press, 1962).

VARIATIONS ON A CUTTING PLANE METHOD FOR SOLVING CONCAVE MINIMIZATION PROBLEMS WITH LINEAR CONSTRAINTS

A. Victor Cabot

Indiana University

ABSTRACT

A cutting plane method for solving concave minimization problems with linear constraints has been advanced by Tui. The principle behind this cutting plane has been applied to integer programming by Balas, Young, Glover, and others under the name of convexity cuts.

This paper relates the question of finiteness of Tui's method to the so-called generalized lattice point problem of mathematical programming and gives a sufficient condition for terminating Tui's method.

The paper then presents several branch-and-bound algorithms for solving concave minimization problems with linear constraints with the Tui cut as the basis for the algorithm. Finally, some computational experience is reported for the fixed-charge transportation problem.

I. INTRODUCTION AND NOTATION

In 1964, Tui [14], discovered an ingenious cutting plane method for solving concave minimization problems with linear constraints. This method has much in common with the recent work in integer programming by Young [15], Balas [1], and Glover [7]. The purpose of this paper is to point out that some of the shortcomings of Tui's original method can be overcome by using recent developments in mathematical programming.

In section II of this paper, the original cutting plane idea of Tui is introduced. In section III the relationship of Tui's cuts and the so-called generalized lattice point problem of linear programming is discussed. A sufficient condition for optimality of the cutting plane algorithm is presented. Section IV presents some branch-and-bound algorithms for concave minimization problems with linear constraints and section V presents computational results of the algorithms of section IV to fixed-charge transportation problems.

In this paper, we consider the mathematical program P: minimize $f(x_1, \ldots, x_{n+m}) = z$ subject to: $(x_1, \ldots, x_{n+m}) \epsilon S$. Here $S = \{(x_1, \ldots, x_{n+m}) \mid \sum_{j=1}^{n+m} a_{ij}x_j = b_i \; i = 1, \ldots, m, x_j \geqslant 0 \; j = 1, \ldots, n+m\}$. We assume that the set S is bounded (this can easily be done if S is "regularized" in the manner of Charnes and Cooper [4]). It is further assumed that every basic feasible solution to P is nondegenerate. More will be said on this matter in section V of the paper. The function f is assumed to be concave over E^{n+m}. If f is a concave function, it makes problem P quite difficult to solve since there may be many local optima. The only theoretical tool which might be able to salvage an otherwise difficult situation is the following easily proven theorem:

THEOREM 1: The global minimum of a concave function f over a bounded convex polyhedral set K is taken on at an extreme point of K.

A proof of this theorem can be found in [4].

This theorem is the basis of Tui's method. Before we present his technique we make one more observation. Every extreme point of the set S can be written in the form

$$x_1=\beta_{10}+\sum_{j=1}^{n}\beta_{1j}(-t_j)\geq 0$$

$$\vdots$$

$$x_m=\beta_{m0}+\sum_{j=1}^{n}\beta_{mj}(-t_j)\geq 0$$

$$x_{m+1}=0+t_1\geq 0$$

$$\vdots$$

$$x_{m+n}=0+t_n\geq 0.$$

The matrix form is given by

$$x=\beta_0+\sum_{j=1}^{n}\beta_j(-t_j)\geq 0.$$

The t_j, $j=1, \ldots, n$ represent the current nonbasic variables. The variables $x_1, \ldots, x_{n+m}$ are of course renumbered so that the first m makes up the basic solution.

II. TUI'S CUTTING PLANE METHOD

Tui's method depends on being able to find locally optimal solutions to problem P. Such local optima may be found by using an adjacent extreme point procedure until an extreme point is reached such that there is no adjacent extreme point with a smaller value of z. The value of the objective function at this locally optimum extreme point, which will be denoted by z_u, gives an upper bound on the value of the optimal solution to P.

Now solve the n one variable optimization problems P_j: maximize t_j subject to: $f(\beta_0+\beta_j(-t_j))\geq z_u$, $0\leq t_j\leq \bar{t}$. Here $\bar{t}$ is some large number. Denote the optimal solution to problem P_j by t_j^*: it can be shown that t_j^* is such that either $t_j^*=\bar{t}$ or $f(\beta_0+\beta_j(-t_j^*))=z_u$. Note that the nondegeneracy assumption guarantees that $t_j^*>0$. The problem P_j can be solved using a search method such as bisection search.

Consider the convex set C generated by the points β_0, $\beta_0+\beta_j(-t_j^*)$ $j=1, \ldots, n$. We now state an important theorem concerning this set.

THEOREM 2: [Glover [7]] For any convex combination y of the points β_0, $\beta_0+\beta_j(-t_j^*)$ $j=1, \ldots, n$ there exists $t_j\geq 0$, $j=1, \ldots, n$, such that $y=\beta_0+\sum_{j=1}^{n}\beta_j(-t_j)$ and $\sum_{j=1}^{n}(1/t_j^*)t_j\leq 1$.

PROOF: Any point y which is a convex combination of the points β_0, $\beta_0+\beta_j(-t_j^*)$ $j=1, \ldots, n$

may be written $y=\lambda_0\beta_0+\sum_{j=1}^{n}\lambda_j[\beta_0+\beta_j(-t_j^*)]$, where $\lambda_0+\sum_{j=1}^{n}\lambda_j=1$ and $\lambda_0,\ \lambda_j\geq 0\ j=1,\ .\ .\ .,\ n.$ Clearly, $\sum_{j=1}^{n}\lambda_j\leq 1$ and equating $t_j=\lambda_j t_j^*\leq t_j^*$ we get $\sum_{j=1}^{n}(1/t_j^*)t_j\leq 1$.

Now consider the set $C\cap S$. Since the function f is concave in E^{n+m} it takes on its global minimum over C at an extreme point of C. Since we already know the value of f at every extreme point of C we know the global minimum of f on C. The value of this solution is z_u. Since the set $C\cap S$ is a subset of C, f must take on a minimum value no less than z_u over $C\cap S$. Since $\beta_0\epsilon C\cap S$ and has $f(\beta_0)=z_u$, β_0 must be the global minimum of f over $C\cap S$.

The set $C\cap S$ is defined by the constraints of S and the additional inequality $\sum_{j=1}^{n}(1/t_j^*)t_j\leq 1$. Since we know the global minimum of f over $C\cap S$ we can discard it from further consideration and explore the remainder of S. We do this by attaching the cut

$$\sum_{j=1}^{n}(1/t_j^*)t_j\geq 1 \tag{1}$$

to the constraints representing the set S and begin looking for another local optimum. When we find one we generate another cut like (1) and continue. The cut (1) is the cut proposed by Tui and the method of continually attaching cuts like (1) is one of the ideas suggested by Tui. It should be pointed out that if lower values of z_u are found as the method proceeds the lowest should be used in problems P_j, $j=1,\ .\ .\ .,\ n$.

Methods similar to the above are being used to generate the so-called "intersection cut" of integer programming. Pioneering papers by Young, Balas, and Glover have shown that most of the cutting plane methods for integer programming are variants of the principle developed by Tui. For most of these integer programming methods, finite convergence can be proved. Tui was unable to present such a proof for his method. Despite this shortcoming we show in the next section that Tui cuts can be related to the so-called generalized lattice point problems of linear programming for which a finite cutting plane algorithm is evidently known. A sufficient condition for Tui cuts obtaining the optimal solution of program P is also given.

III. THE GENERALIZED LATTICE POINT PROBLEM

One form of the generalized lattice point (GLP) problem may be stated as follows: Determine an extreme point of the system

$$\sum_{j=1}^{n+m} a_{ij}x_j=b_i \qquad i=1,\ .\ .\ .,\ m \tag{2}$$

$$x_j\geq 0 \qquad j=1,\ .\ .\ .,\ n+m,$$

which satisfies the constraints

$$\sum_{y=1}^{n+m} d_{kj}x_j \leq f_k \qquad k=1, \ldots, K. \tag{3}$$

This problem has been investigated by Kirby, Love, and Swarup [10] and by Glover and Klingman [8]. In particular, Glover and Klingman have developed a finite cutting plane algorithm for the GLP problem based on a variant of Tui cuts. The relationship to the GLP problem and the method of section II is established by letting the constraints (3) be the Tui cuts of section II and the constraints (2) be the constraints of program P.

It has been shown that the optimum of program P is an extreme point of the set (2). After several Tui cuts have been added, an interesting question is whether all the extreme points of the set (2) have been cut off. If they have, then the best solution to program P found so far is optimal. Solving the GLP problem can determine this fact. If an extreme point of (2) is found which satisfies (3), then this extreme point can be cut off using a Tui cut and the GLP problem can be resolved.

Since the set (2) has only a finite number of extreme points and the GLP problem can be solved in a finite number of steps this gives a finite algorithm for solving program P.

An alternative to solving the GLP problem is to solve a mixed integer formulation of the problem. This is given by finding a feasible mixed integer solution to the constraints:

$$\sum_{j=1}^{n+m} a_{ij}x_j = b_i \qquad i=1, \ldots, m; \tag{4}$$

$$\sum_{j=1}^{n+m} d_{kj}x_j \leq f_k \qquad k=1, \ldots, K; \tag{5}$$

$$x_j - My_j \leq 0 \qquad j=1, \ldots, n+m; \tag{6}$$

$$\sum_{j=1}^{n+m} y_j = m; \tag{7}$$

$$x_j \geq 0 \qquad 0 \leq y_j \leq 1 \qquad j=1, \ldots, n+m, \tag{8}$$

$$y_j \text{ integer.}$$

The value M is taken to be a large positive number.

A mixed integer solution to (4)–(9) will be an extreme point of program P since there will be no more than m positive x_j's by constraints (6) and (7). Solving this problem is of course as difficult as solving the GLP problem, but it has the added advantage that in some cases it can be combined with a bounding procedure for program P to yield a sufficient condition for program P to be solved.

If the function $f(x)$ in program P is separable or quadratic, then it is possible to generate a convex function $g(x)$ such that $g(x) \leq f(x)$ for any solution x satisfying (2) and (3). Methods for doing this are essentially to approximate $f(x)$ by the "closest" linear function as pointed out by Falk and Soland [6] or to generate a "closest" function using linear programming, Cabot and Francis [3]. Suppose

that one had such a linear function $g(x)$ and that the best known solution to program P had a value z_u. Clearly the value z_u is an upper bound on the value of the objective function in the optimal solution of program P. Consider the following linear programming problem:

P1:
$$\text{minimize } g(x)$$
$$\text{subject to (4), (5), (6), (7), (8).}$$

THEOREM 3: If x^* is the optimal solution to P1 and $g(x^*) \geqslant z_u$, then the solution which gave z_u is an optimal solution to program P.

This theorem follows immediately from the property that $g(x) \leqslant f(x)$ for all x satisfying (2) and (3). If $g(x^*) \geqslant z_u$, every solution to (2), (3) must have $f(x) \geqslant z_u$. Of course, a more relaxed problem than P1 can be used in the same manner. An example is

P2:
$$\text{minimize } g(x)$$
$$\text{subject to (2), (3).}$$

We now consider several algorithms for solving P.

IV. BRANCH-AND-BOUND ALGORITHMS FOR SOLVING *P*

In this section we show how Tui cuts can be used to construct algorithms for solving program P. First we present a basic algorithm using only Tui cuts.

a. A Bounding Algorithm

STEP 1. Begin by finding local optimums to P and cutting them off using Tui cuts. As the procedure continues, keep track of the values of f and call the minimum of these values z_u. After several steps derive the function $g(x)$ and solve problem P1 or P2. Go to step 2.

STEP 2. If x^* is the optimal solution to problem P1 and P2 and $g(x^*) \geqslant z_u$, quit. If $g(x^*) < z_u$, add the cut $g(x) \leqslant z_u$ to program P and return to step 1.

This algorithm is similar to one of the original algorithms of Tui. Tui's algorithm attempted to partition the set S by constructing cones over which a solution was known. His goal was to eventually cover the entire set with these cones. The method he used for constructing the cones was based on solving linear programming problems with objective functions generated by the points used to determine his cuts. The difference between the linear programs solved in method IVa and those in Tui's method lies in the construction of the objective functions. Tui constructs several objective functions which serve to approximate f over different regions of S. The objective functions in method IVa attempt to bound f over the entire set S. Zwart [18] has shown that Tui's method can cycle and never solve P.

Finite convergence has not been established for method IVa either. If, however, one solves the GLP problem or the mixed integer equivalent from the previous section, then finite convergence is guaranteed. The power of the algorithm lies in the solution of problem P1 or P2. It is not necessary to exhaust the solution set of program P to obtain optimality.

The weakness of the method occurs at step 1. There is no way of assuming that the optimal

extreme point is generated other than accidentally being found as a local optima at step 1 or step 2. Improvements in the method must occur at that point.

b. Branch-and-Bound Algorithm

This algorithm combines the extreme point ranking principle of Murty [13] and Cabot and Francis [3] for solving P with the previous algorithm. It is assumed that one has a linear function $g(x) \leqslant f(x)$ for all $x \epsilon S$ and some upper bound z_u on the optimal value of $f(x)$ over S.

STEP 1. Start to rank the extreme points of S with respect to increasing values of $g(x)$. Every time one encounters an extreme point of S which is a local optima to P, go to step 2.

STEP 2. It is assumed that there are two formulations of the constraints of problem P, one to use for ranking purposes and one to make cuts on. Cut off the local optimum using a Tui cut on the second formulation. If the current value of $g(x)$ in step 1 is z_l, add the constraint $g(x) \geqslant z_l$ to the second formulation and proceed to step 3.

STEP 3. Derive the function $g'(x) \leqslant f(x)$ for all x satisfying formulation two and solve problem P1 or P2 using $g'(x)$ as the objective function. If x^* solves P1 or P2 $g'(x^*) \geqslant z_u$, quit, the solution giving z_u is optimal, if $g'(x^*) < z_u$, add the cut $g'(x) \leqslant z_u$ to the second formulation. Go to step 4.

STEP 4. Continue ranking extreme points of S according to $g(x)$ until $g(x) > z_u$ or the algorithm terminates at step 3.

This algorithm has several advantages over the first method. The method is finite because S has a finite number of extreme points and in the worst case the method of step 1 would find all of them. Also, the method used in step 1 looks only at extreme points of S. This means it is more likely to generate the optimal solution to P than the method in step 2. A referee has pointed out that method IVb is similar to a technique developed by Young [16] in a cutting plane algorithm for pure integer programming.

It should be mentioned that both IVa and IVb give a sequence of increasing lower bounds z on the value of f at the optimum of P. These bounds are given by the solutions to problems P1 or P2.

c. Incorporation into Existing Branch-and-Bound Methods.

Other algorithms for solving P have been advanced by Falk and Soland and Lawler and Wood [11]. The corporation of Tui cuts into these algorithms may have computational advantages. In particular, the Falk-Soland method is a finite procedure for solving P and its convergence may be speeded by using Tui cuts whenever a local optimum is discovered. This seems to be in the spirit suggested by McCormick [12] by using information gained from observed local optimums to speed branch-and-bound approaches to nonconvex problems.

V. COMPUTATIONAL RESULTS

In this section we present limited computational results using algorithms IVa and parts of IVb. The problems chosen for these experiments were fixed-charge transportation problems due to the availability of a FORTRAN program for ranking extreme points of transportation problems.

The fixed-charge problem is formulated as:

P3: $$\text{minimize } f(x) = \sum_{j=1}^{n} f_j(x_j) = z$$

$$\text{subject to } \sum_{j=1}^{n} a_{ij}x_j = b_i \qquad i = 1, \ldots, m$$

$$x_j \geqslant 0 \qquad j = 1, \ldots, n,$$

where

$$f_j(x_j) = \begin{cases} c_j x_j + d_j & \text{for } x_j > 0 \\ 0 & \text{for } x_j = 0. \end{cases}$$

Each function $f_j(x_j)$ is clearly concave for $x_j \geqslant 0$, but Tui's method demands that it be concave for all x_j. If in problem P3 the coefficients a_{ij} and b_i are integers, then in any basic solution to P3 the value of the basic variables may be written as $k/\det(B)$, where k is an integer and $\det(B)$ is the determinant of the current basis, also an integer. Assuming that we know $\det(B)$ for every basic feasible solution to P3 we can determine an integer M equal to the largest possible value for $\det(B)$. Then the minimum positive value any variable x_j can take on in a basic feasible solution to P3 is $x_j \geqslant 1/M$. The function

$$f_j'(x_j) = \begin{cases} c_j x_j + d_j & \text{for } x_j \geqslant 1/M \\ c_j x_j + M d_j x_j & \text{for } x_j < 1/M \end{cases}$$

can then be substituted for $f_j(x_j)$ in P3. This function takes on the same value as $f_j(x_j)$ at each extreme point solution of P3, but has the added property of being concave over the domain of all real values x_j might take on.

A convex function $g(x)$ approximating $f(x)$ can be determined for each $f_j'(x_j)$ separately depending on the upper and lower bounds on x_j, $l_j \leqslant x_j \leqslant u_j$,

(10)
$$g_j(x_j) = c_j x_j + d_j \qquad l_j \geqslant 1/M$$

$$g_j(x_j) = c_j x_j + d_j - [d_j(1 - Ml_j)/(u_j - l_j)](u_j - x_j)$$

$$0 \leqslant l_j \leqslant 1/M.$$

Problem 1:

Supplies are:	23	38	56	66		
Demands are:	55	54	35	22	9	8
c_{ij}	19	6	12	16	13	24
	5	29	8	19	109	26
	38	17	14	23	27	114
	6	20	2	92	29	42
d_{ij}	4	3	0	6	8	7
	5	16	24	9	11	2
	12	5	10	6	9	43
	8	31	6	12	36	19

The two example problems presented here are taken from Gray [9].

Problem 2:

Supplies are:	45	35	20	15		
Demands are:	35	30	25	15	5	5
c_{ij}	.69	.64	.71	.79	1.70	2.83
	1.01	.75	.88	.59	1.50	2.63
	1.05	1.06	1.08	.64	1.22	2.37
	1.94	1.50	1.56	1.22	1.98	1.98
d_{ij}	11	16	18	17	10	20
	14	17	17	13	15	13
	12	13	20	17	13	15
	16	19	16	11	15	12

The value of M in both of these problems is one. In both problems perturbations can be made so that no solution is degenerate. (See [5], page 314.) Using the function $f_j'(x_j)$ and solving the perturbed problem approximates the original problem. The degree of approximation corresponds to how much the problem was perturbed. In this case the perturbation factor was 0.0001. Each problem was first attacked using the extreme point ranking approach only. The objective function used in the ranking procedure was $g(x)$ in (10) with $l_j = 0, j = 1, \ldots, n$. This corresponds to Balinski's [2] approximate method for transportation problems. For problem 1 the optimal solution was found and confirmed after ranking only two extreme points and using 5 seconds on the CDC–6600 computer.

Problem two was much more difficult. Over 400 extreme points were ranked without confirming optimality. It took 60 seconds to rank the 400 extreme points. It was first thought that the bad results on problem two were due to the numerous extreme points introduced by the perturbation procedure. Further testing on problem one indicated that extreme point ranking is a very sensitive procedure. For many problems no optimal solution could be confirmed. A FORTRAN program of method IVa was then used to solve the problems. The procedure was to add a Tui cut and then do step 2 of the method to check for optimality. For problem 1 two Tui cuts resulted in the feasible region being empty. For problem 2 two Tui cuts resulted in a violation of the upper bound constraint for problem P2 of step 2. Each problem took 12 seconds. A majority of this time was spent determining upper and lower bounds on the variables in step 1 in order to get $g(x)$.

The optimal solution to problem 1 is,

$$x_{12}=14,\ x_{15}=9,\ x_{21}=24,\ x_{24}=6,\ x_{26}=8,\ x_{32}=40,\ x_{34}=16,\ x_{41}=31,\ x_{43}=35,\ z\simeq 1{,}999.$$

The optimal solution to the problem 2 is,

$$x_{11}=20,\ x_{13}=25,\ x_{22}=30,\ x_{26}=5,\ x_{31}=15,\ x_{35}=5,\ x_{43}=15,\ z\simeq 202.$$

It should be noted that the solution to problem 2 is highly degenerate. This seemed to have no effect on the performance of method IVa. One would think that this would lead to shallow cuts in some cases, but none were encountered.

The same two problems were then solved several times with different objective functions. The values of the fixed and variable costs were generated randomly with the ratio of fixed to variable costs varied. In each case five problems were solved. The results are presented in Table 1.

TABLE 1

d/c	Prob. No.	Prob. 1 cut	Prob. 2 cuts
5	1	6^{nf}	3^{nf}
	2	4^{bd}	4^{bd}
	3	2^{bd}	>10
	4	>10	3^{bd}
	5	4^{bd}	5^{nf}
60	1	7^{nf}	$3nf$
	2	>10	5^{nf}
	3	>10	>10
	4	7^{nf}	4^{bd}
	5	6^{nf}	>10
200	1	9^{nf}	2^{bd}
	2	2^{nf}	>10
	3	2^{nf}	9^{nf}
	4	2^{nf}	4^{nf}
	5	2^{nf}	1^{nf}

^{nf}—this implies that the number of cuts indicated completely exhausted the feasible region of the original problem.

^{bd}—this implies that the number of cuts indicated resulted in termination of the algorithm after solving problem P2.

>10—this implies that the algorithm was terminated after 10 Tui cuts were added.

The results of the experiment are mildly encouraging. The algorithm solved 23 of the 30 problems in 9 or less cuts. The fact that problem 2 is degenerate had little effect on the method.

These results seem somewhat contradictory to those obtained earlier by Zwart [17]. The difference could be in the characteristics of the problems solved. Zwart's test problems contained many local optima with objective function values which were close to each other. The problems generated here tended to have few local optimums with widely different objective function values. In most cases the first local optimum found was the global optimum. This tended to make the cuts quite powerful.

REFERENCES

[1] Balas, Egon, "The Intersection Cut—A New Cutting Plane for Integer Programming," Operations Research, *19*, 19–39 (1971).

[2] Balinski, M. L., "Fixed-Cost Transportation Problems," Nav. Res. Log. Quart. *8*, 41–54 (1961).

[3] Cabot, A. V. and R. L. Francis, "Solving Certain Nonconvex Quadratic Minimization Problems by Ranking the Extreme Points," Operations Research Vol. 18, No. 1 (1970).

[4] Charnes, A. and W. W. Cooper, *Management Models and Industrial Applications of Linear Programming* (Wiley, N.Y., 1961).

[5] Dantzig, G. B., *Linear Programming and Extensions* (Princeton University Press, Princeton, N.J., 1963).

[6] Falk, J. and R. M. Soland, "An Algorithm for Separable Nonconvex Programming Problems," Management Science Vol. 15, No. 9 (May 1969).

[7] Glover, Fred, "Convexity Cuts and Cut Search," Operations Research *21*, 123–134 (1973).

[8] Glover, Fred and D. Klingman, "The Generalized Lattice Point Problem," Operations Research *21*, 141–155 (1973).

[9] Gray, P., "Mixed Integer Programming Algorithms for Site Selection and Other Fixed-Charge Problems Having Capacity Constraints," Department of Operations Research Report No. 6, Stanford University, Palo Alto, California (1967).

[10] Kirby, M. J. L., H. R. Love, and Kanti Swarup, "Extreme Point Mathematical Programming," Dalhousie University, (Feb. 1970).

[11] Lawler, E. L. and D. Wood, "Branch-and-Bound Methods: A Survey," Operations Research *18*, 699–719 (1966).

[12] McCormick, G., "Attempts to Calculate Global Solutions of Problems That May Have Local Minima," School of Engineering and Applied Science, The George Washington University, Washington, D.C.

[13] Murty, K., "Solving the Fixed-Charge Problem by Ranking the Extreme Points," Operations Research, Vol. 16, No. 2 (Mar.–Apr., 1968).

[14] Tui, Hoang, "Concave Programming Under Linear Constraints," Soviet Mathematics, 1437–1440 (1964).

[15] Young, R. D., "New Cuts for a Special Class of 0–1 Integer Programs," Research Report in Applied Mathematics and Systems Theory, Rice University, Houston, Texas (Nov. 1968).

[16] Young, R. D., "Finite Pure Integer Programming Algorithms Employing Only Hyperspherically Deduced Cuts," presented at 39th meeting of Operations Research Society of America, Dallas (May 1971).

[17] Zwart, P. B., "Computational Aspects of the Use of Cutting Planes in Global Optimization," Proceedings of the 1971 Annual Conference of the Association of Computing Machinery, Chicago 457–465 (1971).

[18] Zwart, P. B., "Nonlinear Programming: Counter-examples to Global Optimization Algorithms Proposed by Ritter and Tui," Department of Applied Mathematics and Computer Sciences, Washington University, St. Louis, Mo.

OPTIMIZATION IN MIXED-INTEGER SPACE WITH A SINGLE LINEAR BOUND*

Thomas A. Lambe

Department of Industrial Engineering
University of Toronto
Toronto, Ontario

ABSTRACT

The search for an optimal point in a mixed-integer space with a single linear bound may be significantly reduced by a procedure resembling the Lagrangian technique. This procedure uses the coefficients of the linear bound to generate a set of necessary conditions that may eliminate most of the space from further consideration. Enumerative or other techniques can then locate the optimum with greater efficiency. Several methods are presented for applying this theory to separable and quadratic objectives. In the maximization of a separable concave function, the resulting average range of the variables is approximately equal to the maximum (integer) coefficient of the constraint equation.

The division of a one-dimensional resource among several potential applications is a fairly common problem. It arises in the allocation of a large purchase order among a number of vendors [1]. It occurs in the design of equipment redundancy to achieve maximum reliability for a given cost [3]. It appears in the allocation of warehouse stalls to minimize inventory procurement costs, and in the diversification of an investment portfolio to maximize expected return minus risk [4, 6]. The latter two problems are described in this paper.

In all of the previous examples, the optimal decision is to select the quantity x_j of each type of item (j) to maximize some measure $F(x_1, \ldots, x_N)$ of their joint value, where the weighted sum of the quantities cannot exceed a known limit, L. Often the quantities are restricted to non-negative integers. In mathematical terms, the general problem can be formulated as . . .

$$\text{maximize } F(x_1, \ldots, x_N) \tag{1}$$

subject to

$$\sum_{j=1}^{j=N} h_j x_j = L \tag{2}$$

$$x_j \text{ integer for } j = 1, 2, \ldots, P \leq N \tag{3}$$

$$a_j \leq x_j \leq b_j, \tag{4}$$

where h_j positive integer.

---

*Presented in part to the 38th National Meeting of the Operations Research Society of America on October 30, 1970.

Although $F(x_1, \ldots, x_N)$ can be any single-valued function within the constraints, the presentation of the theory in this paper has been simplified by requiring the function to approach minus infinity in any region not satisfying Bounds (4). This procedure does not affect the optimum, nor does it raise any computational difficulties as the discontinuity can be easily recognized. It also does not alter separable nor concave properties of the objective if such initially exist. Although Bounds (4) now are redundant, they are retained to facilitate the application of the theory to specific problems.

The stipulation that coefficients (h_j) are positive integers is not restrictive. If any of these coefficients initially had a fractional component, multiplication of the constraint equation by the (integer) denominator would convert the coefficient to an integer, while preserving the integer properties of the other coefficients. If any of the coefficients (h_j) initially was negative, the replacement of its variable (x_j) by the negative $(y_j = -x_j)$ would produce a positive coefficient. The objective function (1) and bounds 4 must be adjusted for the negative variable. Finally, an inequality constraint may be transformed to a strict equality by the addition of a slack variable with $h_j = 1$.

In general, computational effort to find an optimal solution grows rapidly with the number of variables N and their potential range. To evaluate the objective function for all possible combinations of the values, as would be necessary with a general quadratic objective, may lead to an immense number of calculations. However, in the special case where the objective is the sum of separate functions, one for each variable, dynamic programming greatly reduces the computational burden [5]. When the objective also is piece-wise linear, branch-and-bound techniques may be even more efficient [1, 7]. Finally, if all variables are continuous, Lagrange multiplier techniques may be most effective [2].

The amount of calculation in all of the foregoing cases may be lowered by any method for reducing the region known to contain the optimum. The following observation provides the bases for several such methods.

OBSERVATION: No point $(x_1, \ldots, x_N)$ satisfying Constraints 2 to 4 can be optimum if, for any (i) and $(k \neq i)$,

$$F(x_1, \ldots, x_N) < F(x_1, \ldots, x_i + h_k, \ldots, x_k - h_i, \ldots, x_N).$$

PROOF: The generated point satisfies Constraint 2 because the original point $(x_1, \ldots, x_N)$ satisfies this constraint, and the adjustments cancel each other in Equation (2).

$$h_1x_1 + \ldots + h_ix_i + h_ih_k + \ldots + h_kx_k - h_kh_i + \ldots + h_Nx_N = \sum_{j=1}^{j=N} h_jx_j = L.$$

The generated point also satisfies Constraint 3 because the original point satisfies this constraint, and h_i and h_k are integers, so that

a) if x_i must be integer, $(x_i + h_k)$ also is integer;
b) otherwise x_i and $(x_i + h_k)$ are real numbers;
c) if x_k must be integer, $(x_k - h_i)$ also is integer,
d) otherwise x_k and $(x_k - h_i)$ are real numbers.

Now by definition, no point satisfying the Constraints 2 to 4 can have greater value than the optimum. Furthermore, generated points which do not satisfy Constraint 4 have negative infinite values and hence cannot exceed the optimum. Therefore if the generated point has higher value, $(x_1, \ldots, x_N)$ cannot be optimum. End of proof.

NOTE: Without the stipulation of a negative infinite objective outside of Bounds 4, the foregoing inequality requires other bounds, which make the presentation of the following material more complicated, especially the final section.

The outcome of the foregoing observation is the requirement that an optimal point $(x_1^*, \ldots, x_N^*)$ must satisfy Constraints 2 to 4, plus the $N(N-1)$ constraints defined by the following inequality:

$$(5) \qquad F(x_1, \ldots, x_i+h_k, \ldots, x_k-h_i, \ldots, x_N) \leq F(x_1, \ldots, x_N) \qquad \text{for all } (i, k \neq i).$$

In general, Inequality 5 reduces the number of points that need to be evaluated while searching for an optimum. All regions that do not satisfy this relationship (and Constraints 2 to 4) can be ignored. The ease with which these regions can be recognized, however, depends on the functional form of the objective. Two broad classes are investigated in subsequent sections of this paper.

The effectiveness of Inequality (5) may be improved by a few minor modifications to the coefficients (h_i, h_k). To tighten the constraints, both coefficients should be divided by any common factor that they share. Such an adjustment does not violate Constraints 2 and 3. Also, if the objective never decreases with any increase in a variable having $h_j=1$, it may be advantageous to reduce h_k until a common factor appears, or to increase h_i similarly, or both. In each case, the failure to increase x_j to take up the slack cannot raise the objective. The adjustment reverses with a nonincreasing objective for $h_j=1$. Finally the coefficient h_j of any variable x_j that is not restricted to integers can be made equal to unity by rescaling the variable to $y_j=h_jx_j$.

In the case where x_i and x_k both are real numbers, Inequality (5) is equivalent to a Lagrangian equation. This relationship can be seen by dividing both coefficients (h_i, h_k) by any positive number (n), subtracting $F(x_1, \ldots, x_i, \ldots, x_k-h_i/n, \ldots, x_N)$ from both sides of Inequality (5), and introducing the constant $\lambda_{ik}h_ih_k/n$ between the two. A rearrangement of the terms produces:

$$\begin{aligned}(6) \qquad &[F(x_1, \ldots, x_i+h_k/n, \ldots, x_k-h_i/n, \ldots, x_N)-F(x_1, \ldots, x_i, \ldots, x_k \\ &\quad -h_i/n, \ldots, x_N)]n/h_ih_k \leq \lambda_{ik} \leq [F(x_1, \ldots, x_i, \ldots, x_k, \ldots, x_N) \\ &\quad -F(x_1, \ldots, x_i, \ldots, x_k-h_i/n, \ldots, x_N)]n/h_ih_k.\end{aligned}$$

As "n" approaches infinity, the left and right sides of Expression (6) become the right and left derivatives of the function, if they exist.

$$(7) \qquad \frac{1}{h_i}\frac{\partial^R}{\partial x_i}F(x_1, \ldots, x_N) \leq \lambda_{ik} \leq \frac{1}{h_k}\frac{\partial^L}{\partial x_k}F(x_1, \ldots, x_N).$$

Because Inequality (7) also is valid when the roles of the two variables (i and k) are reversed, the expression becomes a strict equality in the regions where the derivatives also are continuous in these variables. Furthermore in that all of these variables are so linked by Inequality (5), the optimal point must satisfy the resulting Lagrange equations:

$$\frac{1}{h_i}\frac{\partial}{\partial x_i}F(x_1,\ .\ .\ .,x_N)=\frac{1}{h_k}\frac{\partial}{\partial x_k}F(x_1,\ .\ .\ .,x_N)=\ .\ .\ .=\lambda. \tag{8}$$

Separable Objective

When the objective is the sum of separate functions, one for each variable, Inequality (5) becomes a particularly simple expression. All of the terms cancel except the two associated with the coefficients (h_i, h_k). The result is a function of only two variables.

$$F_1(x_1)+\ .\ .\ .+F_i(x_i+h_k)+\ .\ .\ .+F_k(x_k-h_i)+\ .\ .\ .+F_N(x_N)\leq\sum_{j=1}^{j=N}F_j(x_j) \tag{9}$$

$$F_i(x_i+h_k)-F_i(x_i)\leq F_k(x_k)-F_k(x_k-h_i)\qquad\text{for all }(i,\ k\neq i). \tag{10}$$

The number of calculations to find an optimum by dynamic programming is roughly proportional to the sum of the square of the feasible range of each variable. Therefore an obvious avenue for reducing the computational burden is to attempt to narrow the region of search along each variable.

A very simple method for testing whether a given value of a variable (x_g) can be rejected uses upper and lower limits of each variable (x_j) as established by Inequality 10 to form an $(N-1)$-dimensional interval, or hyper-rectangle. Only if such limits exist can x_g be accepted. Furthermore, if the constraint of Equation (2) does not pass through this hyper-rectangle, x_g cannot be a component of the optimum solution because no point $(x_1,\ .\ .\ .,\ x_g,\ .\ .\ .,\ x_N)$ can satisfy both the constraint and Inequality (10). Thus with positive coefficients in the constraint equation, their product with the maximum values of the variables must equal or exceed the constraint constant, L, and their product with the minimum values must equal or be less than the constant.

The foregoing method for reducing the range of each variable has three formal steps. (Shortcuts are given later.)

STEP 1. For any feasible value of x_g, determine upper and lower limits for the optimal value of each variable x_j, and label them $\mathrm{Max}(x_j|x_g)$ and $\mathrm{Min}(x_j|x_g)$, respectively. These limits are established by the following two versions of Inequality (10), plus any previously established limits.

$$F_g(x_g+h_j)-F_g(x_g)\leq F_j(x_j)-F_j(x_j-h_g)$$

$$F_j(x_j+h_g)-F_j(x_j)\leq F_g(x_g)-F_g(x_g-h_j).$$

STEP 2. Reject x_g if a limit does not exist, or if either of the following conditions is not met:

$$\sum_{j=1}^{j=N} h_j \operatorname{Max}(x_j|x_g) \geq L \tag{11}$$

$$\sum_{j=1}^{j=N} h_j \operatorname{Min}(x_j|x_g) \leq L \tag{12}$$

STEP 3. Repeat Steps 1 and 2 for all feasible values of x_g, and for all variables x_1 to x_N.

The preceding method can be illustrated by the allocation of a space 20 yards wide to x_i rows of pallets that are 1 or 2 or 3 yards wide. Each width (i) represents a different production run. Consequently the more rows for a given width, the less frequent is its setup cost. Because 1-yd pallets can occupy any space left by the others, all 20 yards must be allocated when the annual set-up costs (c_i/x_i) are minimized. For purposes of illustration, the problem could be

$$\operatorname{Max} F(x_1, \ldots, x_N),$$

where

$$F(x_1, \ldots, x_N) = -\frac{4}{x_1} - \frac{3}{x_2} - \frac{2}{x_3}, \qquad \text{if all } x_i > 0 \tag{13}$$

subject to

$$= -\infty, \qquad \text{if any } x_i \leq 0,$$

$$x_1 + 2x_2 + 3x_3 = 20 \tag{14}$$

$$x_i \text{ positive integers.} \tag{15}$$

Because the objective equals minus infinity for $x_i \leq 0$, each of the following six relationships from Inequality (10) has a restricted region of applicability. Outside of this region, the right-hand term is replaced by infinity.

$$i=1,\ k=2: \frac{8}{x_1(x_1+2)} \leq \frac{3}{x_2(x_2-1)} \qquad \text{if } x_2 \geq 1 \tag{16}$$

$$i=1,\ k=3: \frac{12}{x_1(x_1+3)} \leq \frac{2}{x_3(x_3-1)} \qquad \text{if } x_3 \geq 1 \tag{17}$$

$$i=2,\ k=1: \frac{3}{x_2(x_2+1)} \leq \frac{8}{x_1(x_1-2)} \qquad \text{if } x_1 \geq 2 \tag{18}$$

$$i=2,\ k=3: \frac{9}{x_2(x_2+3)} \leq \frac{4}{x_3(x_3-2)} \qquad \text{if } x_3 \geq 2 \tag{19}$$

(20) $$i=3,\ k=1:\ \frac{2}{x_3(x_3+1)} \leqslant \frac{12}{x_1(x_1-3)} \quad \text{if } x_1 \geqslant 3$$

(21) $$i=3,\ k=2:\ \frac{4}{x_3(x_3+2)} \leqslant \frac{9}{x_2(x_2-3)} \quad \text{if } x_2 \geqslant 3$$

Table 1 uses Inequalities (16) and (18) to find an upper limit on x_2 for different values of x_1. It starts with the lowest feasible value of x_1. This value establishes 2.67 and infinity in columns 3 and 6 respectively, which become lower and upper bounds for columns 4 and 5, respectively. The highest value of x_2 that satisfies both of these columns is $x_2 = 1$, yielding infinity and 1.50 in columns 4 and 5, respectively. The procedure repeats with successively larger values over the feasible range of x_1. The format of Table 1 also can be used to calculate lower limits for x_2 given x_1.

TABLE 1. *Calculation of* Max $(x_2|x_1)$

1	2	3	4	5	6
x_1	Max x_2	$F_1(x_1+2)-F_1(x_1)$ $=\frac{8}{x_1(x_1+2)}$	$F_2(x_2)-F_2(x_2-1)$ $=\frac{3}{x_2(x_2-1)}$ for $x_2 \geqslant 1$	$F_2(x_2+1)-F_2(x_2)$ $=\frac{3}{x_2(x_2+1)}$	$F_1(x_1)-F_1(x_1-2)$ $=\frac{8}{x_1(x_1-2)}$ for $x_1 \geqslant 2$
1	1	2.67	∞	1.50	∞
2	2	1.00	1.50	0.50	∞
3	2	0.53	1.50	0.50	2.67
4	3	0.33	0.50	0.25	1.00
5	4	0.23	0.25	0.15	0.53
6	4	0.17	0.25	0.15	0.33
7	5	0.13	0.15	0.10	0.23

Table 2 uses the upper and lower limits for x_2 and x_3 to test each value of x_1 against the constraint Equation (14). Columns 2 to 5 give the limits of x_j for a given value of x_1, as calculated by the methods of Table 1. The 6th column shows that the lower limits do not satisfy Inequality (12) when $x_1 \geqslant 7$, and the 7th column shows that the upper limits do not satisfy Inequality (11), when $x_1 \leqslant 4$. No other values of x_1 need be tested because the objective is concave. The region where an optimum can exist therefore has been reduced from the feasible region $1 \leqslant x_1 \leqslant 20$ to $5 \leqslant x_1 \leqslant 6$.

Similar calculations for the other variables restrict the optimum to the region $3 \leqslant x_2 \leqslant 4$ and $x_3 = 2$. The subsequent application of these limits to Table 2 makes $x_1 = 6$. In a similar manner, $x_2 = 4$. Thus the calculation of the optimum is trivial, (6, 4, 2) being the only feasible point.

In the case where a variable (say x_2) is continuous instead of being restricted to integers, the maximum value of x_2 given x_1 in column 2 of Table 1 would not necessarily be an integer. However, for com-

TABLE 2. *Reduction of Range of x_1*

1	2	3	4	5	6	7
x_1	Min $(x_2 \| x_1)$	Max $(x_2 \| x_1)$	Min $(x_3 \| x_1)$	Max $(x_3 \| x_1)$	Σh_j Min $(x_j \| x_1) \leq 20$?	Σh_j Max $(x_j \| x_1) \geq 20$?
1	1	1	1	1	6	6
2	1	2	1	1	7	9
3	1	2	1	2	8	13
4	2	3	1	2	11	16
5	2	4	1	3	12	22
6	3	4	2	3	18	23
7	4	5	2	3	21	26

putational ease, the next larger integer could be used. In a similar manner, the upper and lower bounds in Table 2 could be increased and decreased, respectively, by one unit if x_1 was not restricted to integers.

When the feasible range of each variable is large, and the objective is concave, a simple procedure quickly locates the region where the optimum must exist. The basic principle is to calculate the row in Tables 1 and 2 for a value of x_1 that lies in the middle of a region known to contain the optimum. Initially this point is the center of the feasible range of x_1. If the test value violates Column 6 in Table 2, the optimum cannot lie above the test point because concavity and Equation (10) assure that the maximum values of x_j given x_1 must increase with x_1. Therefore the upper half of the region can be discarded. If the test point violates Column 7, the lower half can be discarded. Repetition of this procedure continues to cut the remaining region in half until the optimal region is found. If the initial range of x_1 is $(10)^6$ units, for example, no more than 20 rows in Tables 1 and 2 are evaluated before both Columns 6 and 7 of Table 2 must be satisfied. Subsequent computation of adjacent rows determines the limits for the location of the optimum, as before. The same procedure of test-and-divide can locate Min $(x_j|x_g)$ and Max $(x_j|x_g)$ in Table 1 if algebraic solutions for them are not available.

Separable concave objectives also have the advantage that any two variables with the same value for their constraint coefficient (h_j) can be combined into one variable by the translation of their lower bounds to the origin, and the generation of a new function from the ordered sequence of the most profitable incremental units of the original pair of variables. Because all of these units have the same coefficient, no advantage is gained by selecting them in any other order. The concave property of the separable objective assures that each unit of the new variable corresponds to a successive unit from one or other of the original two variables.

Quadratic Objective

When the objective (1) is quadratic, Inequality (5) becomes a linear function within restricted regions of the feasible space. Thus for

$$F(x_1, \ldots, x_N) = \sum_{j=1}^{j=N} a_j x_j + \sum_{r,s=1}^{r,s=N} b_{rs} x_r x_s \qquad \text{if } a_j \leq x_j \leq b_j, \forall_j \tag{28}$$

$$= -\infty, \text{ otherwise}$$

where

$$b_{rs} = 0 \qquad \text{if } r > s.$$

Inequality (5) yields for all (i, k $\neq$ i)

$$a_i h_k - a_k h_i + \sum_{s=1}^{s=N} (b_{is} h_k - b_{ks} h_i) x_s + \sum_{r=1}^{r=N} (b_{ri} h_k - b_{rk} h_i) x_r \tag{29}$$

$$+ b_{ii} h_k^2 + b_{kk} h_i^2 - h_i h_k (b_{ik} + b_{ki}) \leq 0, \qquad \text{if } x_k - h_i \geq a_k$$

and

$$x_i + h_k \leq b_i \leq \infty, \text{ otherwise.}$$

Linear programming techniques can quickly locate the region(s) where an optimum can exist. The maximization of a variable subject to the Constraints (29), Equation (2) and Bounds 4 finds an upper limit for a variable, and its minimization finds a lower limit. Repetition of this procedure with each variable forms a hyper-rectangle. (Variables that are restricted to integers are rounded to the nearest enclosed integer.) Evaluation of the feasible points within the hyper-rectangle(s) may proceed with, or without, the use of Constraints (29).

The effects of finite values for Bounds 4 is to require the feasible space to be partitioned into regions where various combinations of the constraints apply. The number of regions, however, may be kept to a minimum by the judicious selection of Constraints (29). The formal procedure has two steps:

STEP 1. Maximize any variable subject to Equation (2), Bounds 4, all Inequalities (29) and the bounds to their region of applicability. If a solution exists, determine upper and lower limits to the enclosing N-dimensional interval by maximizing, and then minimizing each variable.

STEP 2. Repeat Step 1 in other regions of the feasible space by removing relevant members of Inequality 29 and adjusting the bounds of applicability accordingly. Alternatively, all feasible points in these regions can be evaluated.

An application of the foregoing procedure occurs in the allocation of $20,000 among three types of investment which are available in units of 1, 2, and 3 thousand dollars, respectively. If the criterion of choice is to maximize the expected value of a quadratic utility function, the objective becomes a function of the expected value and variance of the uncertain outcome (y) of the total investment.

$$\int (y - ay^2)\, Pdf(y)\, dy = \text{Exp}\,(y) - a[\text{Var}(y) + \{\text{Exp}(y)\}^2].$$

The expected value is the sum of the expected returns from each type of investment (j), each of which is proportional to the number of units x_j that are bought. The variance is the sum of the variances of each investment plus the covariances between investments, which are proportional to the self- and cross-products respectively of x_j. The resulting objective therefore is a quadratic function of x_j. On

the assumption that all of the money is invested, the problem may be illustrated by

$$\text{Max}\, F(x_1, \ldots, x_N) \tag{30}$$

where

$$F(x_1, \ldots, x_N) = [4x_1 + 3x_2 + 2x_3 - x_1^2 + x_1x_2 + x_1x_3 - x_2^2 + x_2x_3 - x_3^2], \qquad \text{if all } x_i \geq 0$$
$$= -\infty, \qquad \text{if any } x_i < 0.$$

subject to

$$x_1 + 2x_2 + 3x_3 = 20 \tag{31}$$

$$x_i \text{ nonnegative integer.}$$

Because the objective equals minus infinity for $x_i < 0$, each of the following six linear inequalities has a restricted region of applicability.

$$i=1, k=2: -5x_1 + 4x_2 + x_3 \leq 2 \qquad \text{if } x_2 \geq 1. \tag{32}$$

$$i=1, k=3: -7x_1 + 2x_2 + 5x_3 \leq 3 \qquad \text{if } x_3 \geq 1. \tag{33}$$

$$i=2, k=1: \quad 5x_1 - 4x_2 - x_3 \leq 12 \qquad \text{if } x_1 \geq 2. \tag{34}$$

$$i=2, k=3: \quad x_1 - 8x_2 + 7x_3 \leq 14 \qquad \text{if } x_3 \geq 2. \tag{35}$$

$$i=3, k=1: \quad 7x_1 - 2x_2 - 5x_3 \leq 23 \qquad \text{if } x_1 \geq 3. \tag{36}$$

$$i=3, k=2: -x_1 + 8x_2 - 7x_3 \leq 24 \qquad \text{if } x_2 \geq 3. \tag{37}$$

The maximization of x_1 subject to Constraints (31) to (37) and their regions of applicability ($x_1 \geq 3$, $x_2 \geq 3$, $x_3 \geq 2$) yields $x_1 = 5.9$. Minimization gives $x_1 = 3.0$. Therefore the optimal integer value of x_1 must lie within $3 \leq x_1 \leq 5$. Similar calculations for the other variables give $3 \leq x_2 \leq 5$ and $2 \leq x_3 \leq 4$. This region of the integer space contains two feasible points, (4, 5, 2) and (5, 3, 3), the latter being optimum.

Inspection of the remaining feasible region finds that it cannot contain an optimum. In particular, no point in $x_1 \leq 2$, $x_2 \geq 3$ can satisfy Constraints (31) and (32). Similarly $x_1 \leq 2$, $x_3 \geq 2$ cannot satisfy Constraints (31) and (33). Furthermore $x_1 \leq 2, x_2 \leq 3, x_3 \leq 2$ does not contain any feasible points. Therefore an optimum cannot exist when $x_1 \leq 2$. In a similar manner, the optimum cannot exist in $x_2 \leq 2$, nor in $x_3 \leq 1$. In general, it is not necessary to check all regions when the objective is concave (negative semi-definite) such as this one. The search should stop once a nonzero N-dimensional interval has been found that does not touch the boundary of an unsearched area. Disjoint intervals are not possible with this type of objective.

Computational Efficiency

This section establishes the effectiveness of the techniques of Table 2 in reducing the space containing the optimum when the (separable) objective is concave. It shows in Inequality (43) that the reduced range for each variable must be less than the sum of the constraint coefficients (h_j) of the other variables. Furthermore, it also shows in Inequality (50) that translation of each variable to place its lower bound at zero reduces the constraint constant, L, to less than twice the product of the maximum coefficient times the sum of all coefficients (h_j). The reduced constant, L', becomes the range of the state variable in the subsequent dynamic programming operation. Furthermore, the derivation of this limit reveals that the average range of the variables usually is less than twice the maximum constraint coefficient. Experience with several test problems indicates that in practice the reduced ranges have approximately one half these limits.

The derivation of an upper bound on the range of a variable reduced by Table 2 first establishes a necessary condition between the coordinates of two points satisfying Inequality (10) when the objective is concave. Thus, in locating a minimum integer value for a variable x_g, Table 2 finds a point $(x_1', \ldots, x_g', \ldots, x_N')$ that satisfies those versions of Inequality (10) for which (i) equals (g). Because $F_g(x_g)$ is concave, any increase (e) in its variable cannot raise the left side of Inequality (10). Therefore for $e > 0$,

$$(40) \qquad F_g(x_g'+h_j+e)-F_g(x_g'+e) \leq F_g(x_g'+h_j)-F_g(x_g') \leq F_j(x_j')-F_j(x_j'-h_g).$$

The methods of Table 2 also find a point $(x_1'', \ldots, x_g'', \ldots, x_N'')$ occurring with the maximum integer value of the variable x_g. This point must satisfy those versions of Inequality 10 for which (k) equals (g). Because $F_j(x_j)$ also is concave, any increase (f) in its variable cannot raise the left side of Inequality (10). Therefore for $f > 0$,

$$(41) \qquad F_j(x_j''+h_g+f)-F_j(x_j''+f) \leq F_j(x_j''+h_g)-F_j(x_j'') \leq F_g(x_g'')-F_g(x_g''-h_j).$$

A comparison of Inequality (40) with (41) reveals that $x_g'+h_j+e=x_g''$ and $x_j''+h_g+f=x_j'$ cannot exist together without causing an inconsistency, unless

$$(42) \qquad F_g(x_g'+h_j+e)-F_g(x_g'+e)=F_g(x_g'+h_j)-F_g(x_g')$$
$$=F_j(x_j')-F_j(x_j'-h_g)=F_j(x_j'-f)-F_j(x_j'-h_g-f).$$

Equation (42) requires that $F_g(x_g)$ be linear over the region x_g' to x_g'', and also $F_j(x_j)$ be linear over the region x_j'' to x_j' with a slope that is h_g/h_j times the slope of $F_g(x_g)$. This unusual case, which produces multiple optima, can be avoided by an incremental adjustment to the slope of one of the components in the objective. In general, therefore, $x_g'+h_j < x_g''$ and $x_j''+h_g < x_j'$ cannot occur together, so that if $(x_g''-x_g') > \underset{j\neq g}{\operatorname{Max}}\, h_j$, then $(x_j'-x_j'') \leq h_g$ for all $j \neq g$.

Inequalities 11 and 12 impose a further restriction on the coordinates of the foregoing two points

$$\sum_{j=1}^{j=N} h_j x'_j \geqslant L \geqslant \sum_{j=1}^{j=N} h_j x''_j.$$

Therefore if $(x''_g - x'_g) > \underset{j \neq g}{\mathrm{Max}}\, h_j$, then

$$(x''_g - x'_g) \leqslant \sum_{j \neq g}^{j=N} h_j (x'_j - x''_j)/h_g \tag{43}$$

$$\leqslant \sum_{j \neq g}^{j=N} h_j.$$

Otherwise,

$$(x''_g - x'_g) \leqslant \underset{j \neq g}{\mathrm{Max}}\, (h_j) \leqslant \sum_{j \neq g}^{j=N} h_j. \tag{44}$$

Inequality (43) therefore is an upper bound on the range of a variable that has been reduced by the methods of Table 2. Furthermore any pair of coefficients (h_g, h_j) that have been divided by a common factor in Table 1 can be so reduced in Inequality (43).

The derivation of an upper limit on the translated constant, L', requires the simultaneous consideration of bounding points for all variables. Therefore, a second subscript must be added to each coordinate to denote the variable being bounded. Thus in Equations (40) and (41), x'_{gg} and x''_{gg} are the resulting lower and upper bounds of the variable x_g, while x'_{jg} and x''_{jg} are the accompanying upper and lower bounds of x_j when x_g equals its lower and upper bounds, respectively.

Because the hyper-space formed by the upper bounds, x''_{gg}, must contain at least one point (the optimum) that satisfies Equation (2), translation of each variable to place its lower bound at zero establishes the following upper limit on L':

$$L' = L - \sum_{j=1}^{j=N} h_j x'_{jj} \leqslant \sum_{j=1}^{j=N} h_j (x''_{jj} - x'_{jj}).$$

Substitution of Inequalities (12) and (11) for x''_{rr} and x'_{ss}, respectively yields, for any (r, s),

$$L' \leqslant \sum_{j \neq r} h_j (x''_{jj} - x''_{jr}) + \sum_{j \neq s} h_j (x'_{js} - x'_{jj}) \tag{45}$$

where

$$x''_{jr} = \mathrm{Min}\ (x_j | x''_{rr}) \text{ in Inequality (12)},$$

and

$$x'_{js} = \mathrm{Max}\ (x_j | x'_{ss}) \text{ in Inequality (11)}.$$

To prove that there is some variable, x_r, for which

$$x''_{jj} - x''_{jr} < \operatorname*{Max}_i h_i, \ \forall_j, \tag{46}$$

it is necessary to arrange the order of the variables according to nonincreasing values of the constraint coefficients, i.e.,

$$h_1 \geqslant h_2 \sim . . . \geqslant h_N.$$

Clearly Inequality (46) holds for $j=r=1$. In general, therefore it must hold for all j less than or equal to some k, where $r \leqslant k$. If $k=N$, the relationship is proven. Otherwise

$$x''_{jj} - x''_{ji} < h_1, \qquad \text{for all } j \leqslant k$$
$$\text{where } i \leqslant k,$$

and

$$x''_{mm} - x''_{mi} \geqslant h_1, \qquad \text{for } m = k+1.$$

To establish that $x''_{jm} \geqslant x''_{ji}$, the assumption that $x''_{ji} - x''_{jm} \equiv p > 0$ is shown to lead to an inconsistency. Thus

$$F_j(x''_{ji} + h_m - p) - F_j(x''_{ji} - p) = F_j(x''_{jm} + h_m) - F_j(x''_{jm})$$

$$\leqslant F_m(x''_{mm}) - F_m(x''_{mm} - h_j), \qquad \text{from Equation (41).}$$

$$\leqslant \frac{F_m(x''_{mi} + h_i) - F_m(x''_{mi})}{h_i/h_j}, \qquad \text{from concavity and } x''_{mm} \geqslant x''_{mi} + h_1.$$

$$\leqslant \frac{F_i(x''_{ii}) - F_i(x''_{ii} - h_m)}{h_i/h_j}, \qquad \text{from Equation (41).} \tag{47}$$

$$\leqslant \frac{F_i(x''_{ii}) - F_i(x''_{ii} - h_j)}{h_i/h_m}, \qquad \text{from concavity and } h_j \geqslant h_m.$$

$$< \frac{F_j(x''_{ji} + h_i - q) - F_j(x''_{ji} - q)}{h_i/h_m}, \quad \text{from Equation (41) to define the minimum value for } q > 0.$$

$$\leqslant F_j(x''_{ji} + h_m - q) - F_j(x''_{ji} - q), \qquad \text{from concavity and } h_i \geqslant h_m. \tag{48}$$

When $j \neq i$, the minimization of x_j given x_i (in Table 2) makes x''_{ji} as small as possible without violating Inequality (41). Consequently, $q=\epsilon$ when x_j is continuous, and $q=1$ when x_j is restricted to integers, unless $x''_{ji}=x'_{jj}$ in which case $x''_{jm} \geqslant x''_{ji}$. However, if $x''_{jm} < x''_{ji}$, $p \geqslant \epsilon$ if x_j is continuous, and $p \geqslant 1$ if x_j

is restricted to integers. Therefore Inequality (48) is inconsistent because of the concavity of F_j (—), and

$$x''_{jj} - x''_{jm} \leq x''_{jj} - x''_{ji} < h_1, \qquad \text{for all } j \neq i \text{ if } j \leq m.$$

When $j=i$, Inequality (47) establishes that $x''_{im} + h_m > x''_{ii}$ except for the unusual (and avoidable) special linear case described for Equation (42). Therefore

$$x''_{jj} - x''_{jm} < h_1, \qquad \text{for all } j \leq n, \qquad \text{where } m \leq n.$$

The foregoing procedure repeats for successively higher subscripts until $n=N$, where $m=r$. Similar procedures establish a variable x_s for which

$$x'_{js} - x'_{jj} < h_1, \qquad \forall j. \tag{49}$$

The substitution of Inequalities (46) and (49) into (45) yields an upper bound on L'.

$$\begin{aligned} L' &\leq \sum_{j=1}^{j=N} h_j (x''_{jj} - x'_{jj}) \\ &< h_1 \sum_{j \neq r} h_j + h_1 \sum_{j \neq s} h_j \\ &< 2h_1 \sum_{j=1}^{j=N} h_j, \qquad \text{where } h_1 = \operatorname*{Max}_j h_j. \end{aligned} \tag{50}$$

Although it is possible for Inequality (50) to be satisfied when the average range of the variables, $(x''_{jj} - x'_{jj})$, is greater than twice the maximum constraint coefficient, in practice the average range is often less. Experience with three examples having 16 variables produced averages that were approximately equal to the maximum constraint coefficient. Also, the variable with maximum range and the value of L' were approximately $1/2$ and $1/4$ of Inequalities (43) and (50), respectively. Computation of the bounds took a few seconds of CPU time on the IBM 360–65, and decreased proportionately with fewer variables. In these examples, the values for h_j were less than 20, L was 10,000, and the concave objectives were either $\sum_j (a_j x_j + b_j x_j^2)$, or $\sum_j (c_j/x_j)$.

REFERENCES

[1] Bracken, J. and G. P. McCormick, *Selected Applications of Nonlinear Programming* (John Wiley, New York, 1968), pp. 28–36.

[2] Everett, H., "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," Operations Research *11*, 399–417 (1963).

[3] Kettelle, J. D., "Least-Cost Allocations of Reliability Investment," Operations Research *10*, 249–265 (1962).

[4] Mao, J. C. and C. E. Sarndal, "A Decision Theory Approach to Portfolio Selection," Management Science *12*, 323–333 (1966).

[5] Nemhauser, G. L., *Introduction to Dynamic Programming* (John Wiley, New York, 1966), pp. 80–84, 101–111.

[6] Paine, N. R., "A Case Study in Mathematical Programming of Portfolio Selections," Applied Statistics *1*, 24–36 (1966).

[7] Wagner, H. M., *Principles of Management Science* (Prentice-Hall, Englewood Cliffs, 1970), pp. 295–299.

ON MIXED INTEGER QUADRATIC PROGRAMS

S. C. Agrawal

Dev Nagri College
Meerut, India

ABSTRACT

This paper provides a method for solving mixed integer quadratic programs with the help of cutting-plane technique.

1. INTRODUCTION

This paper is a continuation of papers [1] and [7] where procedures are given for solving a convex quadratic objective function subject to linear constraints with the condition that the variables are integers.

In this paper, we shall study the case in which *only a subset* of the variables is restricted to consist of integers and the objective function is a convex quadratic one with linear constraints. The paper starts on the same line as a paper by Gomory [4] for the mixed integer problem where the objective function is linear.

The problem is to minimize a convex quadratic objective function,

$$Q(x) = Q(x_1, x_2, \ldots \ldots, x_n),$$

subject to

$$\text{(i)} \qquad Ax = b,$$

$$x \geqslant 0,$$

Some components of x are restricted to be integers. Here, A is an $m \times n$ matrix, $m < n$, b is a given m-vector and x is an n-vector. We will assume absence of degeneration for this method. We also leave the cases of unbounded or infeasible solutions and thus the constraints in (i) define a compact and nonempty convex polyhedron of n dimensions.

2. ANALYSIS

This method first treats the above mentioned mixed integer quadratic program as an ordinary quadratic program, that is, the additional condition that some of the variables are integers is not taken into consideration at the first instance. The ordinary quadratic program is solved by Beale's [3] method and we obtain $x \geqslant 0$. If all variables restricted to be integers are indeed integers, then there is clearly the optimum solution to the mixed integer problem. If some x_i which is restricted to be an integer is not found to be an integer, we add a constraint (called a cut) in the same way as Gomory has done

for linear programs [4], [5]. It should be noted that it does not create any difficulty whether the solution obtained by Beale's method is on the boundary or inside the convex polyhedron, as whenever the solution is inside the original convex region, it happens to be the solution on the boundary of the convex set formed by the original constraints and the additional constraints arrived in the process of obtaining the optimum solution of the ordinary quadratic program.

Now, we apply a special parameter t-method of Beale [3] and by its help we are able to find the integral value of x_i (as described in detail in section 3 below). If there are still some other components of x, which are not integers (but are required to be integers), more constraints are added one by one and the process is repeated on every addition until all required components of x are integers.

3. THE ALGORITHM

For the sake of convenience and easy understanding, we use the same terminology with the same notation as applied by Beale [3] for improper (or free) variables (denoted by "u") to distinguish them from the proper (or restricted), nonnegative variables x. The distinction between improper and proper variables should be well-noted. While improper variables are not restricted to nonnegative values, proper variables are restricted to nonnegative values.

On applying Beale's method to the ordinary quadratic program (ignoring the integrality conditions for some of the components of x as stated in section 2 above) the Kuhn-Tucker conditions for the pth test point to be the minimum point for Q, i.e., $\frac{\partial Q^p}{\partial x_h^p} \geqslant 0$ for all proper variables 'x' and $\frac{\partial Q^p}{\partial u_h^p} = 0$ for all free variables 'u', will be satisfied. Suppose, for this minimum solution of Q, the values of all components of x restricted to be integers are not so. We add a constraint to the existing ones (i.e., existing constraints).

Before adding the above-mentioned constraint, it should be clear that all the free dependent variables with a positive or a negative value in the solution, which have already been eliminated from the equations for the proper basic variables and from Q should be disregarded. The equations for the proper variables serve to control these variables during variations of the independent variables and to keep them nonnegative. For a free variable, which is unrestricted and whose value in the final solution is without interest, such a control is unnecessary. Thus, after all substitutions have been made there are only proper variables in the basis [6].

Here, it should be clearly noted that for the purpose of introducing a cut, we treat proper as well as free variables to be nonnegative ones and obtain the cut exactly on the same lines as Gomory did in [4]. This is the essential feature of the algorithm, because then and then alone the Gomory cut can be formed.

Suppose, after applying Beale's method and ignoring the equations having free variables in the basis, we have one of the equations having proper variable x_i for $i \epsilon J$ (the subset of x constrained to be integer valued) as

$$\text{(ii)} \qquad x_i = a_{io} + \Sigma\, a_{ij}(-x_j) + \Sigma\, a_{ik}(-u_k),$$

containing x_j (proper variables) and u_k (free variables) as nonbasic variables. Here, a_{io} is nonintegral.

We are now in a position to introduce a Gomory cut in the same way and with the same reasoning as we do for linear programs. Let the required Gomory cut [2] be

(iii)

$$\begin{aligned} s_1 = -f_{io} &+ \sum_{j\epsilon J} (-f_{ij})(-x_j) \\ &+ \sum_{j\epsilon J} (-f_{ik})(-u_k) \\ &+ \sum_{\{j\notin J,\, a_{ij} \geqq 0\}} (-a_{ij})(-x_j) \\ &+ \sum_{\{j\notin J,\, a_{ij} < 0\}} \left(\frac{f_{io}a_{ij}}{1-f_{io}}\right)(-x_j) \\ &+ \sum_{\{k\notin J,\, a_{ik} \geqq 0\}} (-a_{ik})(-u_k) \\ &+ \sum_{\{k\notin J,\, a_{ik} < 0\}} \left(\frac{f_{io}a_{ik}}{1-f_{io}}\right)(-u_k), \end{aligned}$$

where s_1 is introduced as a basic variable, f_{io} is a nonnegative fractional part of a_{io}, f_{ij} and f_{ik} are the fractional parts of a_{ij} and a_{ik}, respectively, such that $0 < f_{io} < 1$, $0 \leqslant f_{ij} < 1$ and $0 \leqslant f_{ik} < 1$, i.e., or, in other words $a_{io} = [a_{io}] + f_{io}$, where $[a_{io}]$ denotes the largest integer less than or equal to a_{io} and f_{io} is as defined above with similar expressions for a_{ij} and a_{ik}.

The problem as it stands now can be looked upon as if, after solving the given ordinary quadratic program, we want to add a further linear constraint and then to find the solution of the modified problem. Here, the difficulty arises how to proceed further. The case is a simple one when the objective function is linear, because then dual simplex method can serve the purpose. However, in similar circumstances a general parameter t-method given by Beale [3] can be applied.

We add a parameter t to the right-hand side of the Equation (iii). The equation becomes

$$\begin{aligned} s_1 = -f_{io} + t &+ \sum_{j\epsilon J} (-f_{ij})(-x_j) \\ &+ \sum_{j\epsilon J} (-f_{ik})(-u_k) \\ &+ \sum_{\{j\notin J,\, a_{ij} \geqslant 0\}} (-a_{ij})(-x_j) \\ &+ \sum_{\{j\notin J,\, a_{ij} < 0\}} \left(\frac{f_{io}a_{ij}}{1-f_{io}}\right)(-x_j) \end{aligned}$$

$$+\sum_{\{k\notin J,\, a_{ik}\geqslant 0\}} (-a_{ik})(-u_k)$$

$$+\sum_{\{k\notin J,\, a_{ik}<0\}} \left(\frac{f_{io}a_{ik}}{1-f_{io}}\right)(-u_k).$$

Obviously, the value of t, at present, for which our solution is an optimal and feasible one is negative of the constant term $-f_{io}$, i.e., $t=f_{io}$.

The method now gradually diminishes the value of t to zero. If t is less than f_{io}, then $s_1<0$; and so we make s_1 a nonbasic variable. If s_1 contains any nonzero term in any free variable, we make such a variable basic. If we have many such free variables, any one of them may first be chosen. The reason is that before proceeding further all such free variables are to be removed from the nonbasic set [3].

It may happen that more than one basic restricted variable vanishes as it is about to go negative for the same value of t. The case is that of degeneracy. We are here supposing that degeneracy does not occur. However, in case of degeneracy, Beale's process [3] can be applied.

In the end, we reduce t to zero without any basic restricted variable or any partial derivative of Q becoming negative. If there are still some nonintegral variables which are required to be integers, more constraints are added one by one and the process is repeated on each addition until all required variables are integers.

4. PROOF OF FINITENESS

To prove that the algorithm must terminate in a finite number of steps, we begin by using the notations as given in [6], (chapter seven).

Suppose that the ordinary quadratic program Q after ignoring the integrality restrictions has been represented at the kth test point as a function of $(n-m)$ independent variables, denoted by $z_1^k, z_2^k, \ldots, z_{n-m}^k$, which vanish at the test point:

(iv)
$$Q(x)=Q^k(z_1^k,\ldots\ldots,z_{n-m}^k)$$
$$=\left(c_{oo}^k+\sum_{i=1}^{n-m}c_{oi}^k z_i^k\right)\cdot 1+\sum_{h=1}^{n-m}\left(c_{ho}^k+\sum_{i=1}^{n-m}c_{hi}^k z_i^k\right)z_h^k,$$

($c_{hi}^k=c_{ih}^k$ for $i, h=0, 1, \ldots, n-m$; $c_{hh}^k\geqslant 0$ for $h=1,\ldots,n-m$ because of the convexity of Q).

If s of the z_h^k correspond to free variables which have been introduced in the previous iteration steps, then $(n-m-s)$ correspond to proper variables and are, therefore, sign restricted. The remaining $(m+s)$ proper variables x_{vg}, which have positive values at the test point and are in the basis, depend linearly on the independent variables:

(v)
$$x_{vg}=d_{go}^k+\sum_{h=1}^{n-m}d_{gh}^k z_h^k,\quad g=1, 2, \ldots, i, \ldots, m+s.$$

Further, suppose the Kuhn-Tucker conditions

(vi)
$$\frac{\partial Q^k}{\partial z_h^k} \geqslant 0 \qquad \text{for all sign-constrained } z_h^k,$$

$$\frac{\partial Q^k}{\partial z_h^k} = 0 \qquad \text{for all free } z_h^k.$$

for the kth test point to be the minimum point for Q are satisfied. Thus, the value of Q at this stage is the minimum in the feasible domain and the procedure of obtaining it is convergent [6 (chapter seven)].

We now introduce a cut (iii) as given in section 3 above with the understanding that in forming this cut we treat proper as well as free variables to be nonnegative ones. It should be noted that the treatment of free variables as nonnegative ones for the purpose of obtaining a cut does not create any difficulty in convergence of the algorithm for two reasons. Firstly, in forming the cut the coefficient (say, a_{ij}) of the free variable (say, u_1) may be positive or negative. In either case, the cut will be formed easily. The value of u_1 obtained from the above cut having positive or negative constant on the right-hand side will create no difficulty in finding the values of the other basic variables and the value of the objective function. Secondly, the equations having free variables in the basis will be ignored before proceeding further in each iteration. For convenience, we write the cut (iii) as

(vii)
$$s_1^k = -f_{io}^k + \sum_{h=1}^{n-m} l_{ih}^k\, z_h^k,$$

for one of the constrained x_{vi} of (v), where f_{io} and l_{ih}^k $(h = 1, 2, \ldots, n-m)$ denote the same values as described in section 3 for the constraint (iii). We add a parameter t to the right-hand side of (vii) and thus obtain

(viii)
$$s_1^k = -f_{io}^k + t + \sum_{h=1}^{n-m} l_{ih}^k\, z_h^k.$$

It should be noted that the value of this parameter t for which our solution for the ordinary quadratic program is optimal and feasible one is the negative of the constant term $-f_{io}$, i.e., $t = f_{io}$, where $0 < f_{io} < 1$, and thus the value of t is finite and lies between 0 and 1.

We now have the quadratic program (iv) satisfying the conditions (vi) with $m + s + 1$ constraints given by (v) and (viii). We make s_1^k a nonbasic variable. We first choose free variables as basic ones. As the number of these free variables s on the right-hand side is finite, the removal of all such free variables must terminate in a finite number of steps.

Suppose that after s iterations, $m+s$ proper variables x_{vg} $(g=1, 2, \ldots, m+s)$ are given by

(ix)
$$x_{vg} = d_{go}^{k+s} + t_{go}^{k+s} + \sum_{h=1}^{n-m-s+1} d_{gh}^{k+s} \cdot z_h^{k+s};$$

where d_{go}^{k+s} and t_{go}^{k+s} represent terms containing constant and parameter t, respectively, and the term

containing s_1^k is shown as one of z_h^{k+s}. Also,

$$\text{(x)} \qquad Q(x) = \left(c_{oo}^{k+s} + \sum_{i=1}^{n-m-s+1} c_{oi}^{k+s} z_i^{k+s}\right) \cdot 1 + \sum_{h=1}^{n-m-s+1} \left(c_{ho}^{k+s} + \sum_{i=1}^{n-m-s+1} c_{hi}^{k+s} z_i^{k+s}\right) z_h^{k+s};$$

where c_{oo}^{k+s} and c_{ho}^{k+s} represent terms containing constants and the parameter t alone. The case, as it now stands, is similar to that of the ordinary program having $m+s$ constraints given by (ix) and the quadratic program (x). Hence, the optimal point will be obtained after a finite number of steps [6]. For it, we ultimately reduce t to zero. The solution obtained is either the required optimal solution as the components of x restricted to be integers are all integers or we add a cut for one of the variables required to be integer, but is not so.

Let us suppose that after reducing t to zero we have

$$x_{vg} = c'_{go} + \sum_{h=1}^{n-m-s+1} d'_{gh} z'_h \ (g = 1, 2, \ldots, i, \ldots, m+s),$$

and

$$Q(x) = \left(c'_{oo} + \sum_{i=1}^{n-m-s+1} c'_{oi} \cdot z'_i\right) \cdot 1 + \sum_{h=1}^{n-m-s+1} \left(c'_{ho} + \sum_{i=1}^{n-m-s+1} c'_{hi} z'_i\right) z'_h.$$

We are further assuming that one of the c'_{go} (say, c'_{io}) is not integral, while it is required to be an integer. We obtain a cut

$$s'_1 = -f'_{io} + \sum_{h=1}^{n-m-s+1} l'_{ih} z'_h;$$

and on adding t to the right-hand side, we have

$$s'_1 = -f'_{io} + t + \sum_{h=1}^{n-m-s+1} l'_{ih} z'_h.$$

We now establish that after a finite number of cuts have been added the required optimum solution to the problem will be obtained. We assume that some upper bound is known for the value of Q, i.e., if an integer solution exists for some of the variables required to be integers, it is smaller than some known upper bound M (which can be a large positive constant). This is not a very restrictive assumption and it is always satisfied if the convex region defined by (i) is bounded.

Let c'_{io} be the constant term in the source row, and let c''_{io} be its value after a pivot. It is obvious (see [5], chapter 15) that after a pivot step, the constant term in the source row is either decreased to the next smaller integer $[c'_{io}]$ or increased to the next larger integer $[c'_{io}]+1$. In the beginning, we can put all p constraints where the left-hand sides represent integer-valued variables above those constraints which are only required to be nonnegative.

Since c'_{io} $(i = 1, 2, \ldots, p)$ are all required to be integers, the first p rows are possible source rows. The amount of decrease or increase of c'_{io} in the source row is to the next integer, c'_{oo} cannot

increase indefinitely and still be below the assumed upper bound. Therefore, the first row will be used as the source row if c'_{io} is noninteger. Thus, we can move to the second, third, and finally pth row until all c'_{io} $(i=1, 2, \ldots, p)$ are integers.

5. NUMERICAL EXAMPLE

For the sake of simplicity and easy understanding of the method, we shall now solve an example.

EXAMPLE:

$$\text{Minimize } C=183-44x_1-42x_2+8x_1^2-12x_1x_2+17x_2^2$$

subject to

$$2x_1+x_2 \leqslant 10,$$

$$x_1,\ x_2 \geqslant 0,$$

$$x_1 \text{ is required to be an integer.}$$

By introducing a slack variable x_3, the constraints can be written in the form

$$2x_1+x_2+x_3=10,$$

$$x_1 \geqslant 0, \qquad x_2 \geqslant 0, \qquad x_3 \geqslant 0.$$

The problem when solved by Beale's method [3] gives

$$x_1=\frac{19}{5}+\frac{1}{5}\,u_2-\frac{2}{5}\,x_3,$$

$$x_2=\frac{12}{5}-\frac{2}{5}\,u_2-\frac{1}{5}\,x_3,$$

and

$$C=(19+3x_3)+(4u_2)u_2+(3+x_3)x_3;$$

showing that the minimum solution is

$$x_1=\frac{19}{5}, \qquad x_2=\frac{12}{5}, \qquad \text{and } C=19.$$

We now add a Gomory cut for x_1, because x_1 is required to be an integer here. The cut can be written as

$$s_1=-\frac{4}{5}+\left(\frac{\frac{4}{5}\times-\frac{1}{5}}{1-\frac{4}{5}}\right)(-u_2)+\frac{2}{5}\,x_3$$

$$=-\frac{4}{5}+\frac{4}{5}u_2+\frac{2}{5}x_3.$$

Adding a parameter t to the right-hand side of s_1, we have

$$s_1=-\frac{4}{5}+t+\frac{4}{5}u_2+\frac{2}{5}x_3.$$

We now make s_1 as a nonbasic variable in place of u_2. Thus, we have

$$u_2=1-\frac{5}{4}t+\frac{5}{4}s_1-\frac{1}{2}x_3,$$

$$x_1=4-\frac{1}{4}t+\frac{1}{4}s_1-\frac{1}{2}x_3,$$

$$x_2=2+\frac{1}{2}t-\frac{1}{2}s_1,$$

and

$$C=(19+3x_3)+\left\{4\left(1-\frac{5}{4}t+\frac{5}{4}s_1-\frac{1}{2}x_3\right)\right\}\times\left\{1-\frac{5}{4}t+\frac{5}{4}s_1-\frac{1}{2}x_3\right\}+(3+x_3)x_3$$

$$=(19+3x_3)+\{4-5t+5s_1-2x_3\}\times\left\{1-\frac{5}{4}t+\frac{5}{4}s_1-\frac{1}{2}x_3\right\}+(3+x_3)x_3$$

$$=\left\{23-10t+\frac{25}{4}t^2+\left(5-\frac{25}{4}t\right)s_1+\left(1+\frac{5}{2}t\right)x_3\right\}+\left\{5-\frac{25}{4}t+\frac{25}{4}s_1-\frac{5}{2}x_3\right\}s_1+\left\{1+\frac{5}{2}t-\frac{5}{2}s_1+2x_3\right\}x_3.$$

We now reduce t to zero without any basic restricted variable or any partial derivative of c becoming negative, so we have the solution

$$x_1=4, \qquad x_2=2, \qquad \text{and } C=23.$$

This is an optimum integer solution to the given quadratic program. It should be noted in this example that the value of x_2 also comes out to be an integral one, but it is not essential as only x_1 is required to be an integer.

ACKNOWLEDGMENTS

I wish to express my gratefulness to Professor J. N. Kapur, M.A., Ph.D., F.N.I., Vice-Chancellor, Meerut University, Meerut, for his encouragement. Thanks are also due to Prof. Kanti Swarup, Indian Institute of Public Administration, New Delhi, and Dr. H. C. Jain, Defence Science Laboratory, Delhi–6

for discussions during the course of preparation of this paper. I am indebted to the referee whose constructive remarks has improved the content and presentation of this paper.

REFERENCES

[1] Agrawal, S. C., and Kanti Swarup, "On Integer Solutions to Quadratic Programs," Fourth Annual Convention of the Operational Research Society of India, Indian Institute of Technology, Madras.

[2] Balinski, M. L., "Integer Programming: Methods, Uses, Computation," Management Science, *12*(3), 253–313 (Nov. 1965).

[3] Beale, E. M. L., "On Quadratic Programming," Nav. Res. Log. Quart. *6*, 227–243 (1959).

[4] Gomory, R. E., "An Algorithm for the Mixed Integer Problem," RAND Report P–1885 (Feb. 1960).

[5] Hu, T. C., *Integer Programming and Network Flows* (Addison-Wesley, 1969).

[6] Kunzi, H. P., and W. Krelle, *Nonlinear Programming* (Blaisdell Publishing Co., 1966).

[7] Kunzi, H. P., and W. Oettli, "Integer Quadratic Programming." In R. L. Graves and P. Wolfe, *Recent Advances in Mathematical Programming* (McGraw Hill, 1963).

A SADDLE-POINT THEOREM FOR A CLASS OF INFINITE GAMES

John W. Wingate*

McDonnell Douglas Astronautics Company
Huntington Beach, California

ABSTRACT

In this paper, the existence of a saddle point for two-person zero-sum infinite games of a special type is proved. The games have continuous bilinear payoff functions and strategy sets which are convex, noncompact subsets of an infinite-dimensional vector space. The closures of the strategy sets are, however, compact. The payoff functions satisfy conditions which allow the use of dominance arguments to show that points in the closure of a strategy set are dominated by or are strategically equivalent to points in the strategy set itself. Combining the dominance arguments with a well-known existence theorem produces the main result of the paper.

The class of games treated is an extension of a class studied by J. D. Matheson, who obtained explicit solutions for the saddle points by using necessary conditions.

I. INTRODUCTION

J. D. Matheson [4] introduced a class of infinite two-person zero-sum games in order to handle certain finite games he was investigating in connection with missile defense. Missile-defense problems of the type he investigated can be expressed as matrix games with large numbers (typically millions) of pure strategies. By tacitly using the strong assumption that the utility of an outcome was directly proportional to the number of surviving targets, he was able to use Dresher's method for treating separable games [1, ch. 11] to obtain a formulation of the missile defense problems as games with bilinear payoff functions over compact, convex, strategy sets, these sets being the convex hulls of the sets corresponding to the pure strategy sets of the matrix games. Because of their large number of extreme points, these strategy sets are difficult to use. Matheson sought suitable approximations to the strategy sets.

Among the parameters of the games were three integers: T, the number of targets; aT, the number of attacking missiles; and bT, the number of defending interceptor missiles. Matheson observed that the parameter T often did not appear in the solutions. This observation led him to examine the limiting games obtained as T increased without limit. The strategy sets for the limiting games can be defined simply, and are convex subsets of an infinite-dimensional space. Matheson applied necessary conditions for saddle points and showed that exactly one pair of strategies (for which he gave formulae) in each game satisfies the conditions. In conjunction with an existence theorem, this would show that the games have unique saddle points. However, the standard theorems applied to games with bilinear payoff functions played over convex sets do not apply to Matheson's games, for reasons given later in this section. This paper supplies a suitable existence theorem. It has a wider application than to missile-defense games and is accordingly stated for more general payoff functions than these considered by Matheson.

---

*Present address: Mathematical Analysis Division, U.S. Naval Ordnance Laboratory, White Oak, Silver Spring, Md. 20910

Those interested in applying his results should consult his report [4] or the monograph by Eckler and Burr [2, §4.3]. (The latter gives no proofs.)

Some notational conventions used throughout the paper will now be given. I is the unit interval $[0, 1]$. Z is a countably infinite cartesian product of unit intervals, indexed by the nonnegative integers. Thus if $z \epsilon Z$, z is an infinite sequence $(z_0, z_1, z_2, \ldots, z_n, \ldots)$ of numbers in I. Limits of summation are omitted when they are zero and infinity; that is, '$\sum_n$' means the same as '$\sum_{n=0}^{\infty}$'. For any real number c, the set Z_c is defined by the equation:

$$Z_c = \left\{ z \epsilon Z : \sum_n z_n = 1 \quad \text{and} \sum_n n z_n = c \right\}.$$

The class of games to which the chief result applies will be called 'Matheson games,' defined as follows:

DEFINITION 1: A *Matheson kernel* is a function mapping ordered pairs of nonnegative integers into real numbers that is nonincreasing in its first argument, nondecreasing in its second argument, and has a bounded range.

The arguments of Matheson kernels will be indicated by subscripts. Throughout this paper it will be assumed that Matheson kernels have their ranges in I. This is merely a normalizing condition, and can always be arranged by a suitable change of utility unit and origin.

DEFINITION 2: A *Matheson game* is a two-person zero-sum game for which there are positive real numbers a and b and a Matheson kernel p, such that the sets of pure strategies are Z_a and Z_b and the player whose strategy set is Z_b receives the payoff $\sum_{i,j} p_{ij} x_i y_j$ when he chooses the pure strategy y and his opponent chooses the pure strategy x. Since the numbers a and b and the function p uniquely determine and are uniquely determined by a given Matheson game, this game is called the (a, b, p) Matheson game, and a, b and p are the *parameters of the game.*

Matheson himself studied the subclass of Matheson games (as defined here) in which the payoff kernel p is defined by:

(1)
$$p_{ij} = \begin{cases} q_1^i & \text{if } 0 \leq j < i, \\ q_1^j q_0^{(i-j)} & \text{if } 0 \leq j < i, \end{cases}$$

where $0 \leq q_0 \leq q_1 \leq 1$ and $q_1^0 = 1$ if $q_1 = 0$.

For a real number c, the set Z_c can be considered as a convex subset of an infinite-dimensional vector space, the space of all sequences of real numbers. If there is a locally convex Hausdorff topology for this space (or suitable subspace) such that for the Matheson game with parameters a, b, and p the sets Z_a and Z_b are compact and the payoff function is continuous on $Z_a \times Z_b$, then the game has a saddle point and optimal pure strategies. This follows from a theorem (stated in section II) due to Ky Fan [3]. Unfortunately there is, in general, no such topology. (If (1) defines p, there is no such topology unless $q_0 = 1$; that is, unless the payoff is constant. If $q_0 < 1$, there is a strategy y in Z_b such that the payoff maps $Z_a \times \{y\}$ into a noncompact subset of I. Hence either the payoff is not continuous or Z_a is not compact.)

This difficulty is circumvented by introducing a topology which makes Z compact, applying Fan's theorem to the game over the closures of Z_a and Z_b, and showing that any strategy in the closure of Z_a is dominated by or strategically equivalent to a strategy in Z_a itself (and likewise for Z_b). Thus the chief result, proved in the next section, is:

THEOREM: Every Matheson game has a saddle point and optimal pure strategies.

II. PROOF OF THE THEOREM

The proof of the theorem just stated is based on three lemmata concerning the topological and game-theoretic properties of the sets Z_a and Z_b and the extended-real-valued function F defined on $Z \times Z$ by:

$$F(x, y) = \sum_{i,j} p_{ij} x_i y_j.$$

The set Z is the cartesian product of a countably infinite collection of unit intervals. The topological notions used in this section refer to the product topology for Z. (The unit interval, of course, has its usual topology.) In this topology Z is a compact Hausdorff space satisfying the first axiom of countability, so that sequences are adequate for handling all topological arguments and every sequence in Z has a subsequence convergent to an element of Z. A sequence $(z^1, z^2, \ldots, z^n, \ldots)$ converges to z^0 if and only if for each nonnegative integer i the sequence of real numbers $(z_i^1, z_i^2, \ldots, z_i^n, \ldots)$ converges to z_i^0.

According to Tonelli's Theorem, when x and y are in Z, the quantities $\sum_{i,j} p_{ij} x_i y_j$, $\sum_i \sum_j p_{ij} x_i y_j$, and $\sum_j \sum_i p_{ij} x_i y_j$ are all equal; because of this, $F(x, y)$ will usually be expressed by an iterated summation. The range of the payoff function (the restriction of F to $Z_a \times Z_b$) is contained in I, since if $x \epsilon Z_a$ and $y \epsilon Z_b$,

$$0 \leqslant p_{ij} x_i y_j, \leqslant x_i y_j,$$

so that

$$0 \leqslant \sum_i \sum_j p_{ij} x_i y_i \leqslant \sum_i \sum_j x_i y_j = 1.$$

Although $F(x, y)$ may be infinite for some (x, y) in $Z \times Z$, it is bounded (in fact no greater than 1) for (x, y) in $\bar{Z}_a \times \bar{Z}_b$.

LEMMA 1: For each positive real number a, Z_a is not closed, and

$$\bar{Z}_a = \cup \{Z_c : c \leqslant a\}. \tag{2}$$

PROOF: Let W be the set $\cup \{Z_c : c \leqslant a\}$. The proof has two parts: (a) $W \subset \bar{Z}_a$ and (b) $\bar{Z}_a \subset W$.

(a) $W \subset \bar{Z}_a$. Let x be a point of W. Then $\sum_i x_i = 1$ and $0 \leqslant \sum_i i x_i = c \leqslant a$. For each nonnegative integer i let δ^i be the sequence in Z whose only nonzero term is δ_i^i, which is 1. Let m be the largest integer not greater than a. Consider the sequence defined by:

$$c_n = (m + n - a)/(m + n - c), \qquad n = 1, 2, \ldots$$

and

$$x^n = c_n x + (1-c_n)\delta^{m+n}, \qquad n=1, 2, \ldots .$$

Each x^n is in Z_a since $\sum_i x_i^n = 1$ and $\sum_i i x_i^n = a$. Furthermore,

$$\lim_{n\to\infty} x^n = x$$

since for each nonnegative integer i

$$\lim_{n\to\infty} x_i^n = \lim_{n\to\infty} (c_n x_i + (1-c_n)\delta_i^{m+n}) = \lim_{n\to\infty} c_n x_i = x_i.$$

Thus $x \epsilon \bar{Z}_a$, demonstrating that (a) is true.

(b) $\bar{Z}_a \subset W$. Let x be a point of $\bar{Z}_a$. Then some sequence $(x^1, x^2, \ldots, x^n, \ldots)$ of points in Z_a converges to x. It must be shown that

$$\sum_i x_i = 1 \tag{3}$$

and

$$\sum_i i x_i \leqslant a. \tag{4}$$

The inequality (4) follows from Fatou's lemma, since

$$\sum_i i x_i = \sum \lim_{n\to\infty} i x_i^n \leqslant \lim_{n\to\infty} \inf \sum_i i x_i^n = a.$$

The equation (3) is equivalent to

$$\sum_i \lim_{n\to\infty} x_i^n = \lim_{n\to\infty} \sum_i x_i^n, \tag{5}$$

since $\sum_i x_i^n = 1$ for each n. Now (5) will be true if the sequence of partial sums $\sum_{i=0}^{N} x_i^n$ converges uniformly in n. For each n, $\sum_i i x_i^n = a$, so that

$$N \sum_{i=N+1}^{\infty} x_i^n \leqslant \sum_{i=N+1}^{\infty} i x_i^n \leqslant a.$$

Consequently,

$$\sum_{i=0}^{N} x_i^n = 1 - \sum_{i=N+1}^{\infty} x_i^n \geqslant 1 - a/N.$$

It follows that the sequence of partial sums converges to 1 uniformly in n.

The next lemma is proved by showing that the hypotheses of Fan's minimax theorem [3, Theorem 3] hold for the game over $\bar{Z}_a$ and $\bar{Z}_b$. For completeness, this theorem is quoted here.

THEOREM (K. Fan): Let L_1, L_2 be two locally convex topological linear spaces, and K_1, K_2 be two compact convex sets in L_1, L_2, respectively. Let f be a real-valued continuous function on $K_1 \times K_2$. If, for any $x_0 \epsilon K_1$, $y_0 \epsilon K_2$, the sets $\{x \epsilon K_1 | f(x, y_0) = \underset{\xi \epsilon K_1}{\text{Max}}\, f(\xi, y_0)\}$ and $\{y \epsilon K_2 | f(x_0, y) = \underset{\eta \epsilon K_2}{\text{Min}}\, f(x_0, \eta)\}$ are convex, then

$$\underset{x \epsilon K_1}{\text{Max}}\ \underset{y \epsilon K_2}{\text{Min}}\ f(x, y) = \underset{y \epsilon K_2}{\text{Min}}\ \underset{x \epsilon K_1}{\text{Max}}\ f(x, y).$$

(It should be noted that Fan uses the not universally observed convention that topological linear spaces are always Hausdorff spaces.)

LEMMA 2: $\bar{Z}_a$ and $\bar{Z}_b$ are compact and convex: F is bilinear and continuous on $\bar{Z}_a \times \bar{Z}_b$. There is a point (x^*, y^*) in $\bar{Z}_a \times \bar{Z}_b$ such that for each x in $\bar{Z}_a$ and each y in $\bar{Z}_b$

$$F(x^*, y) \leqslant F(x^*, y^*) \leqslant F(x, y^*).$$

PROOF: Assume temporarily that the first two statements of the lemma have been proved. For each real number c and each y in $\bar{Z}_b$, the set $\{x \epsilon \bar{Z}_a : F(x, y) = c\}$ is convex, since F is linear in x and $\bar{Z}_a$ is convex. In particular, $\left\{x \epsilon \bar{Z}_a : F(x, y) = \max\limits_{\xi \epsilon \bar{Z}_a} F(\xi, y)\right\}$ is convex. Likewise, for each x in $\bar{Z}_a$, $\left\{y \epsilon \bar{Z}_b : F(x, y) = \min\limits_{\eta \epsilon \bar{Z}_b} F(x, \eta)\right\}$ is convex. Then, in view of the theorem just cited,

$$\max_{x \epsilon \bar{Z}_a}\ \min_{y \epsilon \bar{Z}_b} F(x, y) = \min_{y \epsilon \bar{Z}_b}\ \max_{x \epsilon \bar{Z}_a} F(x, y).$$

This equation is equivalent to the last statement in the lemma.

$\bar{Z}_a$ and $\bar{Z}_b$ are compact since they are closed subsets of the compact space Z. Establishing convexity is simply a matter of verifying that $\bar{Z}_a$ and $\bar{Z}_b$ satisfy the definition. F is clearly bilinear; its continuity is not so evident. Since $\bar{Z}_a \times \bar{Z}_b$ is compact, F is continuous if and only if it is uniformly continuous, or, equivalently, if and only if for each positive ϵ there exist finite sets of nonnegative integers M and N, and a positive number δ such that

$$|F(x, y) - F(\bar{x}, \bar{y})| < \epsilon$$

whenever (x, y) and $(\bar{x}, \bar{y})$ are in $\bar{Z}_a \times \bar{Z}_b$, $|x_i - \bar{x}_i| < \delta$ for each i in M, and $|y_j - \bar{y}_j| < \delta$ for each j in N.

For any point x of $\bar{Z}_a$ and for any nonnegative m, $m \sum\limits_{i=m}^{\infty} x_i \leqslant \sum\limits_{i=m}^{\infty} i x_i \leqslant a$. Given a positive ϵ, choose an m greater than $6a/\epsilon$. Then

$$\sum_{i=m}^{\infty} x_i < \epsilon/6.$$

Likewise, let n be greater than $6b/\epsilon$. Then $\sum\limits_{j=n}^{\infty} y_j < \epsilon/6$. Let δ be the minimum of $\epsilon/6m$ and $\epsilon/6n$, and

let M be $\{0, 1, \ldots, m-1\}$ and N be $\{0, 1, \ldots, n-1\}$. Suppose that (x, y) and $(\bar{x}, \bar{y})$ are in $\bar{Z}_a \times \bar{Z}_b$, that for each i in M, $|x_i - \bar{x}_i| < \delta$, and that for each j in N, $|y_j - \bar{y}_j| < \delta$. Then

$$
\begin{aligned}
|F(x, y) - F(\bar{x}, \bar{y})| &= \left| \sum_i \sum_j p_{ij} x_i y_j - \sum_i \sum_j p_{ij} \bar{x}_i \bar{y}_j \right| \\
&= \left| \sum_i \sum_j p_{ij} (x_i y_j - \bar{x}_i \bar{y}_j) \right| \\
&\leqslant \sum_i \sum_j |x_i y_j - \bar{x}_i \bar{y}_j| \\
&= \sum_i \sum_j |x_i y_j - x_i \bar{y}_j + x_i \bar{y}_j - \bar{x}_i \bar{y}_j| \\
&\leqslant \sum_i x_i \sum_j |y_j - \bar{y}_j| + \sum_i |x_i - \bar{x}_i| \sum_j \bar{y}_j \\
&= \sum_i |y_j - \bar{y}_j| + \sum_i |x_i - \bar{x}_i| \\
&= \sum_{j=0}^{n-1} |y_j - \bar{y}_j| + \sum_{j=n}^{\infty} |y_j - \bar{y}_j| + \sum_{i=0}^{m-1} |x_i - \bar{x}_i| + \sum_{i=m}^{\infty} |x_i - \bar{x}_i| \\
&< n\delta + \sum_{j=n}^{\infty} (y_j + \bar{y}_j) + m\delta + \sum_{i=m}^{\infty} (x_i + \bar{x}_i) \\
&< \epsilon/6 + \epsilon/6 + \epsilon/6 + \epsilon/6 + \epsilon/6 + \epsilon/6 = \epsilon.
\end{aligned}
$$

Thus,

$$|F(x, y) - F(\bar{x}, \bar{y})| < \epsilon,$$

and F is uniformly continuous on $\bar{Z}_a \times \bar{Z}_b$.

LEMMA 3: If $x \in \bar{Z}_a$ there is a point $\bar{x}$ in Z_a such that for each y in $\bar{Z}_b$

$$F(\bar{x}, y) \leqslant F(x, y), \tag{6a}$$

and if $y \in \bar{Z}_b$ there is a point $\bar{y}$ in Z_b such that for each x in $\bar{Z}_a$

$$F(x, y) \leqslant F(x, \bar{y}). \tag{6b}$$

PROOF: Let x be a point of $\bar{Z}_a$. Then $\sum_i i x_i = c \leqslant a$. Let m be a positive integer not less than $a - c$ and let $\tilde{x}$ be given by

$$\tilde{x}_k = 0 \qquad \text{for } k = 0, 1, \ldots, m-1;$$

$$\tilde{x}_{i+m}=x_i \qquad \text{for } i=0,\,1,\,2,\,\ldots.$$

Then

$$\sum_i i\tilde{x}_i=\sum_i (i+m)x_i=\sum_i ix_i+\sum_i mx_i=c+m\geqslant a.$$

Define $\bar{x}$ by

$$\bar{x}=((m-(a-c))x+(a-c)\tilde{x})/m,$$

so that $\bar{x}$ is a convex combination of $\tilde{x}$ and x and $\bar{x}\epsilon Z_a$.

For each y in $\bar{Z}_b$

$$F(\tilde{x},\,y)=\sum_i\sum_j p_{ij}\tilde{x}_iy_j=\sum_i\sum_j p_{i+m,j}x_iy_j\leqslant\sum_i\sum_j p_{ij}x_iy_j=F(x,\,y),$$

because p is nonincreasing in its first argument. Since $F(\cdot,\,y)$ is linear,

$$F(\tilde{x},\,y)\leqslant F(\bar{x},\,y)\leqslant F(x,\,y).$$

This proves (6a); proof of (6b) is similar.

PROOF OF THE THEOREM: By Lemma 2 there exists a point $(x^*,\,y^*)$ in $\bar{Z}_a\times\bar{Z}_b$ such that for each x in $\bar{Z}_a$ and each y in $\bar{Z}_b$

$$F(x^*,\,y)\leqslant F(x^*,\,y^*)\leqslant F(x,\,y^*).$$

Lemma 3 then shows that there exists a point $(\bar{x},\,\bar{y})$ in $Z_a\times Z_b$ such that for x in $\bar{Z}_a$ and each y in $\bar{Z}_b$

$$F(\bar{x},y)\leqslant F(x^*,y)$$

and

$$F(x,y^*)\leqslant F(x,\bar{y}),$$

so that

$$F(\bar{x},\bar{y})=F(x^*,y^*)$$

and

$$F(\bar{x},y)\leqslant F(\bar{x},\bar{y})\leqslant F(x,\bar{y}).$$

Since $Z_a\subset\bar{Z}_a$ and $Z_b\subset\bar{Z}_b$, $(\bar{x},\,\bar{y})$ is a saddle point for the $(a,\,b,\,p)$ Matheson game.

Lemma 3 can be interpreted in terms of Matheson's game with the payoff given by (1) in the following fashion: The numbers a and b are the weapon densities, the average number of weapons per target for the attacker and defender. Consider the game from the attacker's viewpoint. According to Lemma 1 a strategy x in $\bar{Z}_a\sim Z_a$ has a weapon density less than a. A strategy obtained from x by allocating enough additional missiles to the targets to bring the weapon density up to a is at least as good as x, perhaps better, and lies in Z_a. The proof of Lemma 3 is a variant of this argument.

III. EXTENSIONS

Schreiber [5] considered a continuous version of Matheson's game, in which the weapons no longer come in discrete units. An offensive pure strategy is a cumulative distribution function F on the non-negative real numbers such that $\int_0^\infty x dF(x) = a$. A defensive strategy G has mean b.

The payoff is $\int_0^\infty \int_0^\infty p(x, y)\ dF(x)dG(y)$, where p has the monotonicity and boundedness properties of a Matheson kernel.

The method of proof used in section II can be adapted to this game with the result that a saddle point can be shown to exist if p has a continuous extension to $[0, \infty] \times [0, \infty]$. (If p is given by (1) with i and j now real variables, and if $q_0 = 0$ or $q_1 = 1$, p does not have such a continuous extension. Thus the existence theorem cannot be applied to the perfect weapon cases of Schreiber's game.) In the proof, the cumulative distribution functions are considered as elements of the unit sphere of NBV$[0, \infty]$, which is a weak*-compact set by Alaoğlu's Theorem and the representation of $C^*[0, \infty]$ as NBV$[0, \infty]$.

It seems reasonable that dominance arguments of the type used in Lemma 3 could be combined with topological arguments to produce still further existence theorems.

IV. ACKNOWLEDGEMENTS

Sidney Moglewer and Melvin Dresher gave generously of their time to discuss this work with me, and I would like to thank them for their valuable criticisms and suggestions.

REFERENCES

[1] Dresher, M., *Games of Strategy: Theory and Applications* (Prentice-Hall, Englewood Cliffs, N.J., 1961).

[2] Eckler, A. Ross and Stefan A. Burr, "*Mathematical Models of Target Coverage and Missile Allocation*," Military Operations Research Society, Alexandria, Va. (1972).

[3] Fan, K., "Fixed Point and Minimax Theorems in Locally Convex Topological Linear Spaces," Proc. Nat. Acad. Sci. U.S.A. *38*, 121–126 (1952).

[4] Matheson, J. D. et al., "Preferential Strategies with Imperfect Weapons," Report AR 67–1, Analytical Services Inc., Falls Church, Va. (Apr. 1967).

[5] Schreiber, T. S., "Defense Models XIV, Optimum Strategies in a Battle Involving Many Identical Targets," Paper 23, Lambda Corp., Arlington, Va. (Nov. 1968).

ON NASH SUBSETS OF BIMATRIX GAMES

C. B. Millham

Department of Computer Science and the Environmental Research Center
Washington State University

ABSTRACT

This work considers a class of bimatrix games to which some well-known structure theorems of 0-sum matrix games can be made to generalize. It is additionally shown how to construct such games and how to generate the equilibrium points defining a given game as a member of that class.

INTRODUCTION

A bimatrix game is defined by a pair (A, B) of real mxn matrices together with the Cartesian product $X \times Y$ of all m-dimensional probability vectors X and all n-dimensional probability vectors Y. A point (x', y') in $X \times Y$ is an equilibrium point of the game (A, B) if $x'Ay'^{\top} \geqq xAy'^{\top}$ for all $x \epsilon X$ and $x'By'^{\top} \geqslant x'By^{\top}$ for all $y \epsilon Y$.

Previous work on such games has included the following: Kuhn [7] defined an extreme equilibrium strategy in a particular way and showed that all extreme equilibrium strategies can be obtained by examining square submatrices of the payoff matrices. Mills [12] showed that a pair (x, y) of mixed strategies is an equilibrium point if and only if there exist scalars α, β such that (x, y, α, β) solves:

$$\text{maximize } xAy + xBy - \alpha - \beta$$

subject to

$$Ay \leqslant \alpha l, \qquad xB \leqslant \beta l, \qquad x \geqslant \phi_m, \qquad y \geqslant \phi_n, \qquad \sum_{i=1}^{m} x_i = \sum_{j=1}^{n} y_j = 1,$$

where l is a row or column vector each of whose entries is 1. Mangasarian [13] considered the same problem, and pointed to the use of the algorithm of Balinski [1] for finding all vertices of a polyhedral convex set in solving it. Lemke and Howson [8] showed, by an algebraic argument, that an equilibrium point lies on a path joining a sequence of adjacent extreme points of a certain convex polyhedron, and gave an algorithm which terminates either in an equilibrium point or in an unbounded edge of the polyhedral set. Raghavan [14] studied the properties of equilibrium points in completely mixed games.

The algebraic properties of 0-sum matrix games were explored thoroughly by Gale and Sherman [4] and by Bohnenblust, Karlin, and Shapley [2]; some of these properties were shown in Ref. [10] to generalize to a class of bimatrix games. Construction of a bimatrix game with given equilibrium points was considered in Ref. [11], which also contained a partial result on the old problem of Nash solvability, Ref. [9], in bimatrix games. This work deals with a class of equilibrium points associated with many matrix pairs (A, B) or with submatrices. A structure is presented, it is shown how to construct a

game (A, B) with the characteristics in question, and a procedure is given by which all equilibrium points of this particular class for which A and B are the essential matrices can be found.

In the following, the ith row of $A(B)$ will be $A_i\cdot(B_i\cdot)$, and the jth column of $A(B)$ will be $A\cdot_j(B\cdot_j)$. P_1 will be the row player, taking his payoff from A, and P_2 will be the column player.

I. RANKS OF ESSENTIAL SUBMATRICES

DEFINITION 1: A Nash-solvable bimatrix game Ref. [9], is a bimatrix game such that, if (x, y) and (x', y') are equilibrium points for (A, B), so are (x, y') and (x', y). The set of all equilibrium points of (A, B) is called its *solution set.*

DEFINITION 2: A Nash subset for a game (A, B) is a set $S=\{(x, y)\}$ of equilibrium points for (A, B) such that, if (x, y) and (x', y') both belong to S, then so do (x, y') and (x', y).

S is thus a collection of equilibrium points for (A, B) that, when interchanged, are still equilibrium points for (A, B). We concern ourselves with Nash subsets, S having more than one member.

DEFINITION 3: Let $S=X_1xY_1\subset X\times Y$ be a Nash subset for (A, B); let, for all $x\epsilon X_1$, $M_1(x)=\{i|X_i>0\}$ and for all $y\epsilon Y_1$, let $N_1(y)=\{j|y_j>0\}$. Let $M_1(S)=\bigcup_{x\epsilon X_1} M_1(x)$, let $N_1(S)=\bigcup_{y\epsilon Y_1} N_1(y)$. ($M_1(S)$ and $N_1(S)$ then index the pure strategies that are essential for S.)

Let $A_1(S)$ be the submatrix of A whose elements (a_{ij}) are indexed by $M_1(S)$ and $N_1(S)$, in their natural order, and similarly for $B_1(S)$. $A_1(S)$ and $B_1(S)$ are then the *essential submatrices* for S.

We assume now that X_1 and Y_1 are the maximal strategy sets, defining a maximal set S, for which $A_1(S)$ and $B_1(S)$ are essential. We denote the cardinality of $M_1(S)$ by $m_1(S)$, and the cardinality of $N_1(S)$ by $n_1(S)$; also, for simplicity, we will hereafter write M_1 for $M_1(S)$, N_1 for $N_1(S)$, n_1 for $n_1(S)$, m_1 for $m_1(S)$, A_1 for $A_1(S)$, and B_1 for $B_1(S)$. The rank of $A_1(S)$ will be $r(A_1)$ and similarly, the rank of $B_1(S)$ is $r(B_1)$. We will sometimes use $S(A, B)$ to indicate the maximal Nash subset S for which A and B are the essential matrices.

DEFINITION 4: A strategy x will be said to be completely mixed with respect to a subset M' of $M=\{1, \ldots m\}$ if $M_1(x)=M'$; that is, if $x_i>0$ if and only if $i\epsilon M'$.

We note that if S is nonempty there exists a pair $(x, y)\epsilon S$ such that x is completely mixed with respect to M_1 and y is completely mixed with respect to N_1: take for x any convex combination of the members of X_1 in which all members of X_1 appear with a nonzero coefficient, and similarly for y.

Assume with no loss of generality that A and B have strictly positive elements.

THEOREM 1: Let A_1 have rank d, let $[Y_1]$ be the linear space spanned by the members of Y_1. Let the matrix $(A\cdot j), j\epsilon N_1$, also have rank d. Then dimension $[Y_1]=n_1-d+1$.

PROOF: Let $y\epsilon Y_1$ be completely mixed with respect to N_1. Delete the 0-components of y, so that $y\epsilon E^{n_1}$ and similarly for the other members of $[Y_1]$ and $[X_1]$; there is no loss of generality. Let $z_1, z_2, \ldots z_{n_1-d}$ be a basis for the nullspace of A_1, regarded as a linear transformation from E^{n_1} to E^{m_1}, and let $\lambda_i>0$ be such that $w_i'=y+\lambda_i z_i$ has all nonnegative components; normalize w_i', $i=1, \ldots n_1-d$, so the normalized result w_i is a strategy. Then $y, w_1, \ldots w_{n_1-d}$ are n_1-d+1 linearly independent strategies and it is easy to see that $w_1, \ldots w_{n_1-d}$ are each in equilibrium with $x\epsilon X_1$. Hence dimension $[Y_1]\geqq n_1-d+1$. On the other hand, suppose $y_1, \ldots y_{n_1-d+2}$ are n_1-d+2 linearly independent

members of Y_1. Let α_1 be the payoff to P_1 when P_2 plays y_1, that is, let

$$A_1 y_1 = \begin{bmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_1 \end{bmatrix}.$$

Since A_1 has all positive components, all payoffs are positive and there are positive scalars $\beta_2, \ldots \beta_{n_1-d+2}$, such that

$$A_1(\beta_i y_i) = \begin{bmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_1 \end{bmatrix} \qquad \text{for } i=2, \ldots n_1-d+2.$$

Then there are n_1-d+2 linearly independent solutions to

$$A_1 y = \begin{bmatrix} \alpha_1 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_1 \end{bmatrix},$$

which contradict the assumption that A_1 has rank d. Since the remaining rows of the matrix (A_{1j}), $j \epsilon N_1$, impose no added restriction, the proof is complete.

Similarly, one can show that if $r(B_1)=e$, and the rank of the matrix $(B_i\cdot)$ $i \epsilon M_1$ is e, then dimension $[X_1]=m_1-e+1$. One can thus state the following corollary:

COROLLARY 1: Let E_1 be the space of unit vectors (essential pure strategies) spanning X_1, let E_2 be the space of unit vectors spanning Y_1. Then dimension E_1- dimension $[X_1]=$ dimension E_2- dimension $[Y_1]-r(A_1)+r(B_1)$.

This generalizes Theorem 1 and its corollary of [4] and Theorem 2 of [10]. It appears that extension to a more general class of bimatrix games is unlikely.

II. BOUNDING FACES AND SUPERFLUOUS STRATEGIES

This section addresses the relationship between the bounding faces of a particular cone and superfluous pure strategies. The following definitions are standard and well-known [1, 2, 4, 6], but are included for completeness. A subset C of a finite-dimensional linear space is a cone if, given x_1, x_2 in C and $\lambda_1 \geqq 0$, $\lambda_2 \geqq 0$, $\lambda_1 x_1 + \lambda_2 x_2 \epsilon C$. The dimension of C is that of the smallest linear space containing C; $C^\perp = \{v \mid (v, x)=0 \text{ for all } x \epsilon C\}$, where (v, x) is the inner product of v and x. The polar cone C^* of C is given by $C^* = \{v \mid (v, x) \geqq 0 \text{ for all } x \epsilon C\}$. If $y \epsilon C^*$, $(y)^\perp$ is a supporting hyperplane of C, and $(y)^\perp \cap C$ is a face of C, where $(y)=\{\lambda y \mid \lambda \geqslant 0\}$. If C has dimension p and $(y)^\perp \cap C$ has dimension $p-1$, $(y)^*$ is an extreme halfspace of C and $(y)^\perp \cap C$ is a bounding face (facet) of C.

Let $E^{n+} = \{x \epsilon E^n \mid x_i \geqq 0, i=1, \ldots n\}$, let $I=(1, 1, \ldots 1) \epsilon E^n$, and $I_i = (0, 0, \ldots 1, \ldots 0)$ be the ith unit vector in E^n. Let Y_1 be as before, with its deleted 0's restored, so that $Y_1 \subset E^n$, and similarly, let $X_1 \subset E^m$.

Let $\begin{pmatrix} Y_1 \\ z \end{pmatrix} \subset E^{n+1+}$ be the set of all $(y_1, \ldots y_n, z)$, with $(y_1, \ldots y_n)^{\mathsf{T}} \epsilon Y_1$, and $(A_i \cdot, y) = z$, $i \epsilon M_1$, so that z is the payoff to P_1 when he plays pure strategy i, $i \epsilon M_1$, and P_2 plays $y = (y_1, \ldots y_n)^{\mathsf{T}}$. Drop the assumption $\sum_j y_j = 1$, to replace a polytope by a cone.

LEMMA 1:

$$\begin{bmatrix} Y_1 \\ z \end{bmatrix} = \bigcap_{i \epsilon M_1} (A_i \cdot, -1)^{\perp} \bigcap_{i \notin M_1} (-A_i \cdot, 1)^* \bigcap_{j \epsilon N_1} (I_j)^* \bigcap_{j \notin N_1} I_j^{\perp}.$$

The proof follows from the fact that $(A_i \cdot, y) = z$, $i \epsilon M_1$, for all $(y, z)^{\mathsf{T}} \epsilon \begin{pmatrix} Y_1 \\ z \end{pmatrix}$, and $(A_i \cdot, y) \leqq z$, $i \notin M_1$, and from the implicit assumption $z > 0$, which follows from the fact that $A > 0$.

Letting $\begin{pmatrix} X_1 \\ w \end{pmatrix} \subset E^{m+1+}$ be the corresponding strategy-payoff vector set for P_1, one can similarly show that

$$\begin{pmatrix} X_1 \\ w \end{pmatrix} = \bigcap_{j \epsilon N_1} (B \cdot_j, -1)^{\perp} \bigcap_{j \notin N_1} (-B \cdot_j, 1)^* \bigcap_{i \epsilon M_1} (I_i)^* \bigcap_{i \notin M_1} (I_i)^{\perp}.$$

A bounding face of $\begin{pmatrix} Y_1 \\ z \end{pmatrix}$ is interior if it does not lie in a proper subface of E^{n+1+} [4].

THEOREM 2: Let f_1 be the number of interior bounding faces of $\begin{pmatrix} X_1 \\ w \end{pmatrix}$, f_2 be the number of interior bounding faces of $\begin{pmatrix} Y_1 \\ z \end{pmatrix}$ and let $s_1 = m - m_1$ be the number of pure strategies for P_1 that are not essential for S; similarly, let $s_2 = n - n_1$ be the number of pure strategies for P_2 that are not essential for S. Then $f_1 \leqq s_2$ and $f_2 \leqq s_1$.

PROOF: Let F be an interior bounding face of $\begin{pmatrix} Y_1 \\ z \end{pmatrix}$. by definition, F is not of the form $I_i^{\perp} \cap \begin{pmatrix} Y_1 \\ z \end{pmatrix}$ and must therefore be of the form $(A_i \cdot, -1)^{\perp} \cap \begin{pmatrix} Y_1 \\ z \end{pmatrix}$. Then $i \notin M_1$, for if $i \epsilon M_1$, $(A_i \cdot, -1)^{\perp} \cap \begin{pmatrix} Y_1 \\ z \end{pmatrix} = \begin{pmatrix} Y_1 \\ z \end{pmatrix}$, and F is not proper. Hence $i \notin M_1$, and i is not essential for S, so that $f_2 \leqq s_1$.

This is also a generalization of similar results in [4] and [10].

III. CONSTRUCTION OF NASH SUBSETS

This section is devoted to a process for constructing a game with a predetermined Nash subset S. Related material appears in [11]. We assume for simplicity that the essential submatrices are to be $m \times n$ and that $M_1 = \{1, \ldots m\}$, $N_1 = \{1, \ldots n\}$, and, for simplicity, we assume that transposes will be clear from the context.

THEOREM 3: Let $(y^i, z^i) \epsilon E^{n+1+}$, $\sum_{j=1}^{n} y_j^i = 1$, $\{y^i\}$ linearly independent, $i = 1, \ldots r$, $z_i > 0$, let $(x^k, w^k) \epsilon E^{m+1+}$, $\sum_{i=1}^{m} x_i^k = 1$, $\{x^k\}$ linearly independent, $k = 1, \ldots s$, $w^k > 0$. Regard the matrices

$$\begin{bmatrix} y^1 \\ \cdot \\ \cdot \\ \cdot \\ y^r \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x^1 \\ \cdot \\ \cdot \\ \cdot \\ x^s \end{bmatrix},$$

respectively, as linear transformations form E^n to E^r and E^m to E^s. If the nullspaces of these transformations are nontrivial, there is a game (A, B) with a Nash subset $S=S'xS''$, where $S'=\left\{x \,|\, x=\sum_{i=1}^{s} \lambda_i x^i,\right.$ $\left.\lambda_i \geqq 0, \sum_{1}^{r} \lambda_i = 1\right\}$ and $S''=\left\{y \,|\, y=\sum_{j=1}^{r} \delta_j y^j, \delta_j \geqq 0, \sum_{i}^{r} \delta_j = 1\right\}$, and such that, if $(x, y) \epsilon S$ with payoffs α, β to the respective players and $x=\sum \lambda_i x^i$, $y=\sum \delta_j y^j$, then $\alpha=\sum \delta_i z^i$, $\beta=\sum \lambda_j w^j$.

PROOF: Let $(a_1, \ldots a_\mu)$ be a basis for the nullspace of

$$\begin{bmatrix} y^1 \\ \cdot \\ \cdot \\ \cdot \\ y^r \end{bmatrix},$$

let $(b_1, \ldots b_\nu)$ be a basis for the nullspace of

$$\begin{bmatrix} x^1 \\ \cdot \\ \cdot \\ \cdot \\ x^s \end{bmatrix}.$$

For $i=1, \ldots r$, let a_i' be such that $(a_i', y^i)=z_i$, $(a_i', y^j)=0$, $i \neq j$; this is possible by the linear independence of the $\{y^j\}$; similarly so for b_i', $i=1, \ldots s$. For $k=1, \ldots m$, let

$$A_k \cdot = \sum_{i=1}^{\mu} \lambda_{ik} a_i + \sum_{i=1}^{r} a_i', \lambda_{ik}$$

arbitrary. Similarly, let

$$B \cdot_j = \sum_{i=1}^{\nu} \delta_{ij} b_i + \sum_{i=1}^{s} b_i', j=1, \ldots n.$$

It follows that

$$A = \begin{bmatrix} A_1 \cdot \\ \cdot \\ \cdot \\ \cdot \\ A_m \cdot \end{bmatrix}, \qquad B=(B_{\cdot 1}, \ldots B_{\cdot n})$$

define a game in which $S''=\left\{y|y=\sum_{i=1}^{r}\delta_i y^i,\ \delta_i\geqq 0,\ \sum_i \delta_i=1\right\}$ and $S'=\left\{x|x=\sum_{i=1}^{s}\sigma_i x^i,\ \sigma_i\geqq 0\ \sum_i \sigma_i=1\right\}$ define a Nash subset $S=S'xS''$ for (A, B).

IV. DETERMINING NASH SUBSETS

This section is devoted to a procedure for generating all the equilibrium points in the Nash subset S for which a given pair (A_1, B_1) of submatrices is essential. The material draws heavily on a well-known paper by Davis [3] and on a manuscript by Shephard [15] which has received widespread attention among workers in convexity and related areas. For the sake of completeness some of this work is abstracted and presented here. The initial reference is Davis [3].

Let $a_1 \ldots a_r$ belong to E^n. A positive combination $\sum_{i=1}^{r}\lambda_i a_i$ of $a_1, \ldots a_r$ is a linear combination with $\lambda_i\geqq 0$. If all $\lambda_i>0$ and all $a_i\neq\phi$, the combination is strictly positive. An equation giving ϕ as a positive combination of some subset of $\{a_1, \ldots a_r\}$ is a positive *relation* among the members of that subset. The relation is strictly positive if no vector in that subset is the 0-vector ϕ and no coefficient is 0.

A polyhedral convex cone is defined by the set $\sum_{i=1}^{r}\lambda_i a_i$, of all positive combinations of the set $\{a_1, \ldots a_r\}$. The vectors $a_1, \ldots a_r$ are positively dependent if any a_i is a positive combination of the others; otherwise they are positively independent. A frame of a cone C is a positively independent set $\{a_1, \ldots a_r\}\subset C$ that positively spans C. If C is a linear space, the frame is said to be a positive basis for C. The cone C is pointed if C contains no nontrivial linear space; it can be shown [5] that the frame of a pointed cone is unique.

If $a_1, \ldots a_r$ is a positive basis for a linear space S, it can be shown [3] that there is a strictly positive relation $\phi=\sum_{i=1}^{r}\lambda_i a_i$ among $a_1, \ldots a_r$. If the dimension of S is $r-1$, $a_1, \ldots a_r$ is called a *minimal* positive basis, S is said to be minimal, and the positive relation is unique.

It is shown in [3] that all positive relations among $a_1, \ldots a_r$ can be obtained by taking positive combinations of the strictly positive relations associated with minimal bases found among $\{a_1, \ldots a_r\}$.

In the following it will be again best to omit formal indications of the transposes of vectors. The meaning will always be clear.

Consider the "augmented" game matrices $(A_1, -I_{m_1})$ and $(B_1^{\mathrm{T}}, -I_{n_1})^{\mathrm{T}}$, where $I_{m_1}\epsilon E^{m_1}$, $I_{n_1}\epsilon E^{n_1}$, are vectors whose components are all 1. Assume that A_1 and B_1 are the game matrices A and B. Then the following is clear.

THEOREM 4: (x, y) is an equilibrium point in $S(A, B)$ if and only if, for some $\alpha>0$, $\beta>0$, (y, α) and (x, β) are, respectively, positive relations among $(A\cdot_1, \ldots A\cdot_n, -I_m)$ and $(B_1\cdot, B_2\cdot, \ldots B_m\cdot, -I_n)$.

The proof is simple and is omitted. The implication is that if one can find all positive bases contained, e.g., among $A\cdot_1, \ldots A\cdot_n, -I_m$, and similarly for $B_1\cdot \ldots B_m\cdot, -I_n$, one will be able to get all equilibrium points in S, and further, that the problem reduces to finding all positive relations associated with the minimal subspaces spanned by positively independent subsets of $A\cdot_1, \ldots A\cdot_n$, $-I_m$, and similarly for $B_1\cdot, \ldots B_m\cdot, -I_n$. This is facilitated by the results of Shephard [15], which are cited in greater detail than those of Davis but still without proof.

Let $X=\{x_1 \ldots x_r\}$ be a set of points in E^n. If the cone S spanned by the members of X is not E^n, that cone S has supporting hyperplanes; if H is such a supporting hyperplane, $H \cap S$ is a *face* of S. If the dimension of the linear hull of X is m, faces of dimension $m-1$ are *facets* of S.

Let $L(X)$ be the set of all linear dependencies of $X: L(X)=\left\{\lambda_1, \ldots \lambda_r \,\middle|\, \sum_{i=1}^{r} \lambda_i x_i=\phi\right\}$. Assume the linear hull of X is E^n. Then $L(X)$ has dimension $r-n$; if $\{a_i\}=\{(\alpha_{i1}, \ldots \alpha_{ir})\}, i=1, \ldots r-n$, is a basis for $L(X)$, write

$$M_L(X)=\begin{bmatrix} \alpha_{11} \cdots & \alpha_{1r} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \alpha_{r-n,1} \cdots & \alpha_{r-n,r} \end{bmatrix}.$$

Letting

$$x_i'=\begin{bmatrix} \alpha_{1i} \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{r-n,i} \end{bmatrix}$$

be the ith column of $M_L(X)$, the set $X'=\{x_1' \ldots x_r'\} \subset E_{r-n}$ is a set of *representatives* of $\{x_1, \ldots x_r\}$, and x_i' is a representative of x_i.

Write pos X for the positive hull of X. If $Y \subset X$ is such that pos Y is a face of pos X and $Y=X \cap$ pos Y, then the set $X-Y$ is a *coface* of X.

THEOREM 5 (Shephard): A set $Y \subset X$ is a coface of X if and only if ϕ is in the relative interior of the convex hull of Y'.

COROLLARY 1 (Shephard): A set $X \subset E^n$ positively spans E^n if and only if ϕ does not belong to the convex hull of X'.

THEOREM 6 (Shephard): A set X is a positive basis for E^n if and only if ϕ does not belong to the convex hull of X' and there is no hyperplane through ϕ strictly separating one of the points x_i' from $X'-\{x_i'\}$.

Now, let X positively span E^n, let H be any hyperplane (there is always one) not containing ϕ such that $H \cap$ pos X' is an $r-n-1$-polytope P'. For each $x_i' \epsilon X'$ consider the ray $r(x_i')=\{\lambda_i' x_i' \mid \lambda > 0\}$, and let this ray meet P' at x_i''. The set $X''=\{x_1'', \ldots x_r''\}$ is a *diagram* of X.

THEOREM 7 (Shephard): Let X be a positive basis in E^n, X'' a diagram for X, and F' any facet of the polytope formed by all convex combinations of X''. Then the subset $X''-(F' \cap X'')$ of X'' corresponds to points of X which form a basis for a minimal subspace of E^n with respect to X.

This last result makes it possible to find the equilibrium points in the Nash-subset S by making it possible to pick out the minimal subspaces of $(A\cdot_1, \ldots A\cdot_n, -I_m)$ and $(B_1\cdot, \ldots B_m\cdot, -I_n)$, respectively. When the minimal subspaces have been found, the unique strictly positive relations associated with these minimal subspaces can be found in a straightforward manner. The other equilibrium points in S can then be generated by taking all positive combinations of these minimal positive relations.

PROCEDURE: Reduce, by row operations, $(A, -I_m)$ and $\begin{pmatrix} B \\ -I_n \end{pmatrix}^T$ to row-echelon form or until a basis for the nullspaces of $(A, -I_m)$ and $\begin{pmatrix} B \\ -I_n \end{pmatrix}^T$ can be found. Form the systems Y', X' of representatives of $(A, -I_m)$ and $\begin{pmatrix} B \\ -I_n \end{pmatrix}$, respectively. If any representative $y_j' \epsilon Y'$ is ϕ, exclude column $A_{\cdot j}$ (set $y_j=0$ in any strategy), because $\{A_{\cdot 1}, \ldots A_{\cdot n}, -I_m\}$ do not positively span the space they linearly span if $A_{\cdot j}$ is included. If, after this ϕ is in the convex hull of the remaining set $\{A_{i\cdot}, -I\}$, remove vectors $A_{\cdot i}$ until ϕ is no longer positively expressible in terms of the remaining vectors $\{y_j'\}$. Retain the $A_{\cdot i}$ thus removed for later consideration. If any $A_{\cdot i}$ remain, the remaining $A_{\cdot i}$, together with $-I_m$, positively span the space they linearly span, and a hyperplane not through the origin can be found such that a ray derived from each of the remaining y_j' will intersect the hyperplane at a point y_j''. The polytope obtained by taking all convex combinations of the $\{y_i''\}$ can then be examined for its facets. The coface of any facet will define a minimal positive basis. The unique strictly positive relations among these minimal positive faces will positively span all positive relations among the columns of $(A, -I_m)$, and similarly for $\begin{pmatrix} B \\ -I_n \end{pmatrix}^T$. If x and y are non-ϕ positive relations among the rows of $\begin{pmatrix} B \\ -I_n \end{pmatrix}$ and columns of $(A, -I_m)$, respectively, and are scaled to sum to 1, (x, y) is an equilibrium point of (A, B).

Any excluded $A_{\cdot j}$ whose representative y_j'' is not ϕ must now be reconsidered, some other $A_{\cdot j}$, now excluded so as to avoid a positive representation of ϕ and the process repeated, and similarly for any excluded $B_{j\cdot}$ with a representative $x_j'' \neq \phi$.

If A_1, B_1 are in fact proper submatrices of A and B, respectively, the situation is more complicated and the process will not necessarily work.

EXAMPLE:

$$A=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 2 \\ 1 & 2 & 1 & 4 \end{pmatrix}, \qquad B=\begin{pmatrix} 1 & 2 & 1 & 3 \\ 4 & 3 & 2 & 1 \\ 3 & 1 & 2 & 2 \end{pmatrix}.$$

The "augmented" matrices are

$$A'=\begin{pmatrix} 1 & 2 & 3 & 4 & -1 \\ 2 & 1 & 3 & 2 & -1 \\ 1 & 2 & 1 & 4 & -1 \end{pmatrix}, \qquad B'=\begin{pmatrix} 1 & 2 & 1 & 3 \\ 4 & 3 & 2 & 1 \\ 3 & 1 & 2 & 2 \\ -1 & -1 & -1 & -1 \end{pmatrix}.$$

When A' is reduced to row-echelon form one obtains

$$R=\begin{pmatrix} 1 & 0 & 1 & 0 & -1/3 \\ 0 & 1 & 0 & 2 & -1/3 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

a basis for the nullspace of which is $(1, 1, 0, 0, 3)$, $(0, 1, 0, -1/2, 0)$. The system of representatives

of A' is

$$w=\begin{pmatrix}1\\0\end{pmatrix},\quad \begin{pmatrix}1\\1\end{pmatrix},\quad \begin{pmatrix}0\\0\end{pmatrix},\quad \begin{pmatrix}0\\-1/2\end{pmatrix},\quad \begin{pmatrix}3\\0\end{pmatrix};$$

since $w_3=\phi$, $A\cdot_3$ is excluded from consideration, and one notes that the remaining representatives

$$\begin{pmatrix}1\\0\end{pmatrix},\quad \begin{pmatrix}1\\1\end{pmatrix},\quad \begin{pmatrix}0\\-1/2\end{pmatrix},\quad \begin{pmatrix}3\\0\end{pmatrix}$$

do not convexly span ϕ. One then obtains $1(1, 1, 0, 0, 3)-1/2(0, 1, 0, -1/2, 0)=(1, 1/2, 0, 1/4, 3)$, a positive relation (strict, among $A'\cdot_1, A'\cdot_2, A'\cdot_4, -I$), and notes that the coefficients 1 and $-1/2$ define a system of hyperplanes not all through the origin. Choosing one, say $z_1-1/2z_2=2$, find the intercepts of the rays $\lambda_i w_i$, $i\neq 3$, on the hyperplane to obtain the diagram $\{w_i'\}$.

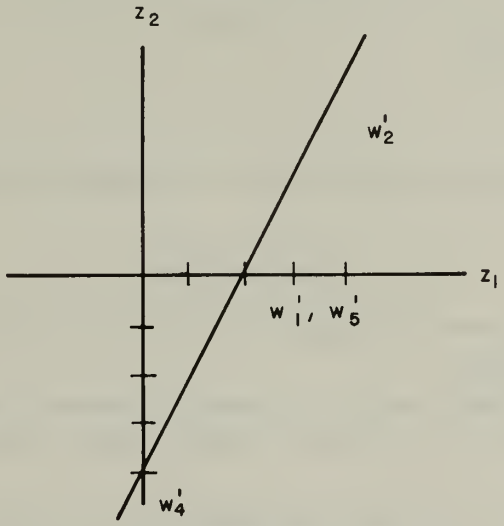

FIGURE 1

The facets of the polytope conv $\{w_1', w_2', w_4', w_5'\}$ are w_4' and w_2'; one coface is $w_1'w_2'w_5'$, the other is $w_4'w_1'w_5'$. Both correspond to minimal positive bases: $T_1=\{A\cdot_1, A\cdot_2, A\cdot_5\}$, $T_2=\{A\cdot_4, A\cdot_1, A\cdot_5\}$. The strictly positive relations associated with T_1 and T_2, respectively, scaled, are $y'=(1/2, 1/2, 0, 0, 3/2)$ and $y^2=(2/3, 0, 0, 1/3, 2)$.

Since $y_3'=0$, one can exclude $B\cdot_3$ and consider

$$B^{*T}=\begin{pmatrix}1 & 4 & 3 & -1\\ 2 & 3 & 1 & -1\\ 3 & 1 & 2 & -1\end{pmatrix}$$

Row-echelon form is

$$R'=\begin{pmatrix}1 & 0 & 0 & -5/20\\ 0 & 1 & 0 & -3/20\\ 0 & 0 & 1 & -1/20\end{pmatrix},$$

from which a strictly positive relation, scaled to sum to 1, is $(^5/_9, {}^3/_9, {}^1/_9, {}^{20}/_9) = x$. One varifies that $(x, B\cdot_3) = {}^{13}/_9 < {}^{20}/_9$, and (x, y^1), (x, y^2) are equilibrium points for (A, B), along with (x, y), $y = \sum_{i=1}^{2} \lambda_i y^i$, $\lambda_i \geqq 0$, $\sum \lambda_i = 1$. The essential submatrices for $S = \{x\}x\left(\sum_{i=1}^{2} \lambda_i y^i,\ \lambda_i \geqq 0,\ \sum \lambda_i = 1\right)$ are

$$A^1 = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 1 & 2 \\ 1 & 2 & 4 \end{pmatrix}, \qquad B^1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}.$$

If one considers instead the game (A, C), for

$$C = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 4 & 3 & 2 & 1 \\ 3 & 1 & 2 & 2 \end{pmatrix},$$

the process fails because $(x, B\cdot_3) > {}^{20}/_9$. Thus the game (A, C) has no Nash subset S in which all pure strategies for P_1 are active.

Acknowledgment

The contributions and suggestions of G. H. Heuer are gratefully acknowledged.

REFERENCES

[1] Balinski, M. L., "An Algorithm for Finding all Vertices of Convex Polyhedral Sets," J. Soc. Indust. Appl. Math. *9*, 72–88 (1961).

[2] Bohnenblust, H. F., S. Karlin, and L. S. Shapley, "Solutions of Discrete, Two-Person Games. Contributions to the Theory of Games," Annals of Math. Studies 24, (Princeton University Press, 1950).

[3] Davis, Chandler, "Theory of Positive Linear Independence," Amer. J. Math., *76*, 733–746 (1954).

[4] Gale, D., and S. Sherman, "Solutions of Finite Two-Person Games," in Contributions to the Theory of Games, Annals of Math. Studies 24 (Princeton University Press, 1950).

[5] Gerstenhaber, M., "Theory of Convex Polyhedral Cones," chap. XVIII of Cowles Comm. Mono. 13, *Activity Analysis of Production and Allocation*, Koopmans ed. (Wiley, New York, 1951).

[6] Goldman, A. J. and A. W. Tucker, "Polyhedral Convex Cones" in *Linear Inequalities and Related Systems*, Annals of Math. Studies 38 (Princeton University Press, 1956).

[7] Kuhn, H. W., "An Algorithm for Equilibrium Points in Bimatrix Games," Proc. N.A.S. *47*, 1656–1662 (1961).

[8] Lemke, C. E. and J. T. Howson, Jr., "Equilibrium Points of Bimatrix Games," J. Soc. Indust. Appl. Math. *12*, 413–423 (1964).

[9] Luce, R. D. and H. Raiffa, *Games and Decisions* (John Wiley and Sons, N.Y., 1957).

[10] Millham, C. B., "On the Structure of Equilibrium Points in Bimatrix Games," SIAM Review *10*, 447–449 (1968).

[11] Millham, C. B., "On Constructing Bimatrix Games with Special Properties," Nav. Res. Log. Quart., *19*, 709–714 1972.

[12] Mills, H., "Equilibrium Points in Finite Games," J. Soc. Indust. Appl. Math. *8*, 397–402 (1960).

[13] Mangasarian, O. L., "Equilibrium Points of Bimatrix Games," J. Soc. Indust. Appl. Math. *12*, 778–780 (1964).

[14] Raghavan, T. E. S., "Completely Mixed Strategies in Bimatrix Games," J. London Math. Soc. *2*, 709–712 (1970).

[15] Shephard, G. C., "Diagrams for Positive Bases," J. London Math. Soc. *4*, 165–175 (1972).

THE OPTIMAL SIZE OF A STORAGE FACILITY

Joel Levy

National Bureau of Standards

ABSTRACT

The appropriate size for a piece of fixed capital equipment (measured in units of capacity) depends on the anticipated demand for its services and on its cost. Using several models developed in the study of optimal inventory policy we derive the contribution to cost reduction that additional storage space makes under each of these models. Comparison of the sum of the discounted benefits (i.e., reduced operating cost) with construction costs for additional storage space then yields the optimal size of the storage facility.

0. INTRODUCTION

The theory that is used to investigate optimal inventory policy has principally been devoted to the study of the stockage policies to be followed so that expected operating costs of a firm shall be minimized. With variations for the structure of the model used, the operating costs consist mainly of storage costs, order costs and shortage costs. Different models allow different forms for the cost functions, i.e., constant, linear, convex or more general. Alternative lag structures between the time an order is placed and the time of its arrival have also been considered. But in most previous studies variations in the size of the storage facility are not considered.*

In the present paper to derive the optimal size of a storage facility we have narrowed the scope of the study to some of the elementary models used in the development of optimal inventory theory. For these models we show that a relatively simple expression of the cost parameters in the model and of the distribution function of demand yields the present value of additional units of storage capacity, i.e., the reduced cost of operating the inventory system discounted to the present. Furthermore it turns out that for these models the present value of storage capacity is a differentiable convex function. Therefore, when construction costs as a function of capacity can be expressed by a differentiable convex function computing the optimal size of the storage facility (given the data mentioned above) is a direct calculation.

1. DETERMINISTIC MODEL WITH TRANSACTIONS COST

One of the early models in inventory theory is a deterministic model in which demand and cost conditions are assumed stationary.† In this model it is assumed that in addition to the price paid for the goods ordered there is a procurement cost for each order which is independent of the size of the order. Storage costs are supposed constant per unit of goods stored per unit time, and sales are assumed to be uniform over time.

---

*An exception to the general statement that the proper size for a storage facility has not been investigated is the set of notes by Ghosal [3], which reports the work of Hurst [4]. There is no overlap between the results reported by Ghosal and the material presented here.

†See, for instance, References [1] or [4].

Under these conditions if we denote unit storage costs per unit time by h, procurement cost per order by K, annual sales by S, and the quantity bought in each procurement by X, the total of procurement and storage costs per year is (as a function of the quantity bought in each procurement):

$$W(X)=\frac{X}{2}h+K\frac{S}{X}.$$

Differentiating with respect to X, we see that for X positive this function is minimized at

$$X^*=\sqrt{\frac{2KS}{h}}.$$

Consider a facility in which the unit storage costs per unit time are h up to the quantity $\bar{X}$ and no more than $\bar{X}$ can be stocked in the facility at any time. If $\bar{X}$ is greater than X^* the limited capacity of the storage facility does not prevent one from using the optimal inventory policy. If $\bar{X}$ is not greater than X^* the best inventory policy satisfying the capacity restriction is to reorder to full capacity each time the stock on hand is exhausted. The total of procurement and storage costs per year of this policy is

$$W(\bar{X})=\frac{\bar{X}}{2}h+K\frac{S}{\bar{X}}.$$

The additional yearly cost due to capacity limitation is

$$\Delta(\bar{X}, X^*)=W(\bar{X})-W(X^*)$$

$$=\frac{h}{2}(\bar{X}-X^*)+KS\left(\frac{1}{\bar{X}}-\frac{1}{X^*}\right).$$

Suppose:

(a) All expenses occurring in a given year are to be discounted at the rate applicable to the start of that year, and (b) an averaging procedure is available to the firm so that the expenses actually incurred in accordance with the timing of the process described are spread uniformly over all years.

Let the rate of interest be denoted by i. If with no change in unit storage costs the capacity of the facility can be expanded from $\bar{X}$ to X^* at a cost of less than $\frac{\Delta(\bar{X}, X^*)}{i}(1+i)$ such an expansion would be profitable. If the cost of expanding storage capacity from $\bar{X}$ to X^* exceeds $\frac{\Delta(\bar{X}, X^*)}{i}(1+i)$ the cost of the expansion exceeds the discounted return from the expenditure.

2. DETERMINISTIC MODEL WITH VARIABLE CAPACITY

Assume the demand and cost conditions of the previous section. Namely, demand is stationary and uniform at a rate of S units per year, unit storage cost per unit time is constant h up to the capacity of the storage facility, and procurement cost per order is a constant amount K independent of the size

of the order. Suppose now the problem is raised of what the capacity of a storage facility to be built should be given that the cost of construction of the facility is an increasing function of its capacity.

It is clear that under the above assumptions there is no reason to build a facility with capacity greater than X^*, where, as in section 1, $X^*=\sqrt{\frac{2KS}{h}}$. This is so since the policy that minimizes annual costs never stores more than X^* units. However, since the costs of building the facility increase with the facility's capacity it is economical to let the capacity remain smaller than X^*.

Take a specific example of a storage facility which as a function of the capacity, Z, for $Z > 0$ costs $V(Z)=C_3Z^3+C_1Z+C_0$ to construct. For the cost functions of inventory that we have assumed the annual cost of replenishing to inventory level X each time the stock is exhausted is

$$W(X)=\frac{X}{2}h+K\frac{S}{X}.$$

Retain the assumptions of section 1 that (a) expenses occurring in a given year are to be discounted at the rate applicable to the start of that year, and (b) the firm is allowed to average its expenses so that they are spread uniformly over all years. Under a rate of interest i the discounted future cost of this policy is $\frac{W(X)\,(1+i)}{i}$. We introduce a symbol for the discounted future cost of the inventory policy of reordering to X from stock level zero and denote it by $D(X)$, so that $D(X)=\frac{W(X)}{i}(1+i)$.

The optimal size of a storage facility is that size which minimizes the sum of the cost of construction plus the discounted future inventory costs when using the facility. Let us denote the sum of these costs as a function of the capacity of the facility, X, by $E(X)$. If the capacity of the facility does not exceed X^* then the discounted future inventory costs are $D(X)$.

Thus we wish to find that value of X, the capacity of the storage facility, that will minimize $E(X)$. By definition X will be positive and we have argued above that the minimizing value of X will not exceed X^*. In the interval between zero and X^* the total cost function is given by

$$E(X)=V(X)+D(X).$$

Each of the functions $V(\cdot)$ and $D(\cdot)$ is convex and differentiable so that the same is true for $E(\cdot)$. Hence the value of X that minimizes $E(\cdot)$ can be found by differentiating E and finding that value of X for which this derivative is 0.

Carrying out the indicated operations, we get

$$E^1(X)=V^1(X)+D^1(X)$$

$$=3C_3X^2+C_1+\frac{h(1+i)}{i2}-\frac{KS(1+i)}{iX^2}=0\,.$$

The following particular cases of this formula may be worth mentioning.

(1) When the cost of expanding capacity is not a feature of the problem this is included in above example by putting $C_3=C_1=0$. The optimal reorder quantity is that obtained for the case where

there is no capacity constraint:

$$X^*=\sqrt{\frac{2KS}{h}}\,.$$

(2) If the cost of constructing the facility is linear in its capacity, this is included in the above example by putting $C_3=0$ and $C_1>0$. In this case the minimizing condition found above yields that the facility capacity and inventory reorder quantity should be

$$X^*=\sqrt{\frac{KS}{C_1\frac{(i)}{1+i}+\frac{h}{2}}}$$

(3) When the cost of constructing the facility is strictly convex in the capacity built and is described by our formula, then $C_3>0$. Finding the solution for the quadratic equation gotten when we put $y=X^2$ and then taking the square root of the positive solution, we obtain that in the indicated situation the optimal size for the storage facility is

$$X^*=\left\{\frac{\sqrt{12C_3KS\frac{1+i}{i}+\left[C_1+\frac{h}{2}\frac{1+i}{i}\right]^2}-\left[C_1+\frac{h}{2}\frac{1+i}{i}\right]}{6C_3}\right\}^{1/2}$$

It may not be amiss to call attention here to the fact that technological improvement resulting in a reduction of construction costs may lead to a smaller size facility being optimal for our problem than would have been the case in the absence of the innovation. While a comparison of alternative total costs would substantiate this assertion, the truth may perhaps more easily be grasped by noting that to minimize total cost we examined marginal costs.

3. STOCHASTIC DEMAND MODEL

An alternative model has often been used in the investigation of inventory policies in the circumstances of stochastic demand.* The assumptions of this model differ from that of the previous sections and are as follows:

a. instead of being a known quantity, it is assumed that the amount demanded in each time period is a random variable the distribution function of which is known;

b. the amounts demanded in different time periods are assumed to be indepencent of one another and to have the same distribution function;

c. an order to replenish inventory (the size of the order may be 0) is placed at the beginning of each time period and the cost of placing the order is proportional to the amount ordered;

d. storage costs are constant per unit stored per period and are charged on the number of units on hand after order;

*This is basically the model discussed in Reference [2], pp. 159–164.

e. a penalty, which is constant per unit, is assessed on the excess of the amount demanded over the amount held in inventory;

f. a constant rate of interest, i, per period is assumed so that one unit of cost incurred in period $t+1$ is equivalent to $\frac{1}{1+i}$ units of cost incurred in period t.

Let $a=\frac{1}{1+i}$, $g(X)$ the frequency function of demand per period, c_1 the cost of ordering a unit of the commodity, c_2 the cost per period of storing a unit of the commodity and p the penalty per unit of the excess of demand over stock on hand. Denote by $f(X)$ the expected sum of future discounted costs of following an optimal inventory policy when the initial stock is X, and by c the sum c_1+c_2. Then $f(X)$ satisfies the functional equation

$$f(X) = \min_{y \geqslant x} \left\{ c_1(y-X) + c_2 y + \int_y^\infty p(z-y)g(z)dz + a\left[f(0)\int_y^\infty g(z)dz + \int_0^y f(y-z)g(z)dz \right] \right\}.$$

From this expression for $f(X)$ one can deduce that the optimal policy is to keep as the stockage objective that quantity y for which

$$\int_0^y g(z)dz = \frac{p-c}{p-ac_1}.$$

Denote the quantity y satisfying this equation by y^*.

For any y let $T(y)$ be the expected cost under the economic model described above of the inventory policy that maintains y as the stockage objective and starts from an initial stock level of zero. Then

$$T(y) = c_1 y + \frac{1}{1-a}\left[c_2 y + \int_0^y p(z-y)g(z)dz + ac_1 y - a\int_0^y c_1(y-z)g(z)dz \right].$$

If the capacity of the storage facility, $\bar{y}$, does not exceed y^* then the policy of restocking to $\bar{y}$ at the start of each period is the one that minimizes expected discounted costs. Starting from a given stock level X the cost of this policy is

$$T(\bar{y}; X) = c_1(\bar{y}-X) + \frac{1}{1-a}\left[c_2\bar{y} + \int_{\bar{y}}^\infty p(z-\bar{y})g(z)dz + ac_1\bar{y} - a\int_0^{\bar{y}} c_1(\bar{y}-z)g(z)dz \right].$$

The cost of the capacity limitation is $T(\bar{y}) - T(y^*)$.

Thus if an expansion of the capacity of the facility from $\bar{y}$ to y^* is possible at a cost less than $T(\bar{y}) - T(y^*)$, such an expansion would be profitable. If the cost of expansion to y^* exceeds $T(\bar{y}) - T(y^*)$ the cost of expansion exceeds the expected return.

4. STOCHASTIC DEMAND MODEL WITH VARIABLE CAPACITY

Assume the demand and cost conditions of the stochastic demand model described in the preceding section. Suppose the problem is now raised of what the capacity of a storage facility to be built should be, given that the cost of construction of the facility is an increasing function of its capacity.

In the economic situation we are dealing with, there is no gain in building a facility with the capacity greater than y^*, where y^* is the amount that would be stocked under the cost and demand conditions assumed were space available. It was mentioned in section 3 that y^* is the amount of inventory that makes the probability of a stockout $\frac{c_2+(1-a)c_1}{p-ac_1}$. However, if construction costs can be reduced by keeping the facility's capacity less than y^*, it may be profitable to do so though this is at the expense of increased expected periodic inventory charges.

As an example suppose as a function of capacity z, for $z>0$, the cost of construction of the facility is:

$$K(z) = K_2 z^2 + k_1 z + k_0.$$

If the capacity of the facility is not greater than y^*, the best inventory policy is to restock to capacity at the start of each period. Denote by y the facility capacity. The expected discounted costs of this policy is

$$T(y) = \frac{1}{1-a}\left[cy + \int_y^\infty p(z-y)g(z)dz - a\int_0^y c_1(y-z)g(z)dz \right].$$

The optimal size of a storage facility is that size which minimizes the sum of the cost of construction plus the discounted future inventory costs when using the facility. Let us denote the sum of these costs as a function of the capacity of the facility, y, by $\sum(y)$. Thus in the range of values relevant to our investigation

$$\sum(y) = K(y) + T(y).$$

Each of the functions $K(\cdot)$ and $T(\cdot)$ is convex and differentiable so that the same is true for $\sum(\cdot)$. Hence the value of y that minimizes $\sum(\cdot)$ can be found by differentiating, and finding that value of y for which this derivative is 0.

Carrying out the indicated operations we get:

$$\begin{aligned}\sum{}'(y) &= K^1(y) + T^1(y)\\ &= 2k_2 y + k_1 + \frac{1}{1-a}\left[c - \int_y^\infty pg(z)dz - a\int_0^y c_1 g(z)dz \right] = 0.\end{aligned}$$

By algebraic manipulation this can be converted to the equivalent equation,

$$2k_2 y + \frac{p-ac_1}{1-a}\int_0^y g(z)dz = \frac{p-c}{1-a} - k_1.$$

5. OPTIMAL FACILITY CAPACITY WHEN DEMAND IS INCREASING

The economic models used earlier in this chapter to specify the situation requiring a facility assumed that demand was stationary. For a situation in which the demand and cost conditions are assumed stationary assessment of whether to expand storage capacity or not, or the optimal size of a facility can be determined by converting the cost of facility expansion into its equivalent uniform annual expenditure. Thus, if as above the rate of interest is denoted by i and the cost of the facility expansion by K, then the amount of periodic expenditure equivalent to a present capital expense of

K is $K \cdot i$ per period. That is instead of, as was done above, translating operating expenses of the future to the present by discounting, the evaluation of the benefit of the storage facility against its cost could be carried out by converting the cost of construction of the facility incurred at present into its equivalent as a periodic expense and adding this to the inventory storage cost. If demand is not stationary an optimal inventory policy cannot, in general, be found by the means that are effective in the inventory problems we discussed above. However, where demand is increasing it has been shown elsewhere that the best inventory policy is the one that is optimal for stationary demand. (See Reference [5].) We assume that the rate of increase in demand is not so rapid that the discounted value of expected amounts demanded is infinite.

To be more precise denote the frequency function of demand in time period j by $g_j(s)$. Suppose that for each j and for every x

$$\int_0^x g_{j+1}(s)ds \leq \int_0^x g_j(s)ds.$$

This condition is what is meant by demand increasing. In actual occurrence the amount demanded in a given period may be smaller than that demanded in a preceding period. Denote by y_j^* a value of y that satisfies

$$\int_0^y g_j(s)ds = \frac{p-c}{p-ac_1},$$

where the symbols p, a, c, c_1, are as defined in section 3. Then an optimal inventory policy is: at each time period j order the amount necessary to attain to the level y_j^*.

Use the symbol $M_j(y)$ to represent the function of y whose value for each nonnegative y is

$$C \cdot y + \int_y^\infty p(s-y)g_j(s)ds - a\int_0^y c_1(y-s)g_j(s)ds.$$

For $y \leq y_{j+1}^*$, $M_j(y)$ can be interpreted as the expected cost to be assigned to time period j when the stock level y is taken as the stockage objective at the start of that period. This interpretation is justified by noting that (1) the first term is the cost of purchasing and storing y units (2) the second term is the expected penalty for being short of amount demanded when y units are acquired, and (3) the third term is the expected value of the material remaining on hand after the jth period, pricing the units at the rate applicable to them. Their value is the contribution they make to a reduction in costs in period $(j+1)$, and so to obtain the corresponding value in period j that cost should be discounted for one period.

If the capacity of the storage facility, $\bar{y}$, does not exceed y_j^* then the policy of restocking to $\bar{y}$ at the start of the jth period is the one that minimizes expected costs to be allocated to that period. The expected cost of the capacity limitation is

$$M_j(\bar{y}) - M_j(y_j^*).$$

Define

$$D_j(\bar{y}, y_j^*) = M_j(\bar{y}_j) - M_j(y_j^*) \qquad \text{for } \bar{y} < y_j^*$$

$$0 \qquad \text{for } \bar{y} \geq y_j^*.$$

Let

$$K(\bar{y}) = \sum_{j=1}^{\infty} a^{j-1} D_j(\bar{y}, y_j^*).$$

$K(\bar{y})$ is the sum of the expected costs attributable to the fact that the storage facility has the limited capacity $\bar{y}$. The costs that it is anticipated will be incurred in any given period are appropriately discounted.

If an expansion of the capacity of the facility from y_1 to y_2 is possible at a cost less than $K(y_1) - K(y_2)$ such an expansion would be profitable. If the cost of that expansion exceeds $K(y_1)-K(y_2)$ the cost of the expansion exceeds the expected return from it.

Let the cost of construction of a facility as a function of its capacity z be denoted by $V(z)$. Suppose this function of z is increasing and convex for $z > 0$.

For each j, $D_j(y, y_j^*)$ is differentiable in y and the derivatives are all bounded by a common constant. Therefore the function $K(y)$ is also differentiable.

Moreover, each of the functions $a^{j-1}D_j(y, y_j^*)$ is convex in y. The partial sums $\sum_{j=1}^{N} a^{j-1}D_j(y, y_j^*)$ and the limit sum $K(y)$ are therefore also convex.

We are seeking the value of capacity that will minimize total expected costs as a function of capacity. Denote the total cost by $E(z)$. It consists of the facility construction cost and the added inventory management cost attributable to the capacity constraint. In terms of the notation previously defined

$$E(z) = V(z) + K(z).$$

Suppose $V(z)$ is differentiable. Convexity of $V(z)$ was one of our hypotheses. The above analysis has shown that for the cost structure of the inventory management model we are using $K(z)$ is convex and differentiable. Under these conditions the minimizing value of z for the sum of these functions can be found by differentiating E and finding a value of z for which the derivative is 0.

BIBLIOGRAPHY

[1] Arrow, Kenneth J., "Historical Background," in *Studies in the Mathematical Theory of Inventory and Production*, edited by Kenneth J. Arrow, Samuel Karlin, and Herbert Scarf (Stanford University Press, 1958).

[2] Bellman, Richard E., *Dynamic Programming* (Princeton University Press, 1957).

[3] Ghosal, A., Some Aspects of Queueing and Storage Systems, (Springer-Verlag, 1970).

[4] Hurst, H. E., "Methods of using Longterm Storage in Reservoirs," Proc. Inst. Civ. Engrs. Paper 6059 (1956).

[5] Levy, Joel, "Optimal Inventory Policy When Demand Is Increasing," *Operations Research* Vol. i, No. 6 (November–December, 1960), pp. 861–863.

[6] Whitin, Thomson M., *The Theory of Inventory Management* (Princeton University Press, 1953).

ON THE EVALUATION OF INVESTMENT PROPOSALS HAVING MULTIPLE ATTRIBUTES

Robert R. Trippi

School of Business Administration
California State University, San Diego

ABSTRACT

A framework is developed for analyzing the likelihood of acceptance of an investment project proposal when objectives are uncertain. The foundation is a utility model of top management's choice process, modified if need be through a Bayesian approach which takes into account any apparent inconsistency in the history of past proposal acceptances and rejections.

The normative uses of limited information for communication and control within the multi-divisional firm is a subject of both theoretical and practical interest. Generalized programming, for example, provides the mathematical framework and economic interpretation for models in which divisional decisions can be expected to achieve overall goals through use of the information represented by internal resource prices. The development of a means of processing limited or seemingly inconsistent information by the division manager in order to formulate investment proposals which are likely to be consistent with global goals is another topic of some interest. Lusk [1] has in a recent article suggested the use of a multiple discriminant function as a means of identifying proposals with a high probability of acceptance. This paper examines a similar problem within the framework of an economic model of behavior.

PROJECT PROPOSALS AND UTILITIES

The division manager has as his information base a history of acceptances and rejections of past project proposals which have been submitted to headquarters by his and possibly other divisions. He wishes to determine whether his current proposal is likely to be accepted. A project is characterized by a vector, each element of which is a measure of some attribute of perceived relevance to top management's decision, as for example expected return, risk, community acceptance, etc.

Top management's decision process is assumed in the simplest case to consist of a comparison of the utility of each project (according to its own utility function $U(x)$, where x is the project attribute vector) with some norm U^*. Projects with utility equal to or in excess of U^* are accepted, while all others are rejected. As an example of how the history of past proposal outcomes might be used by the division to provide information about the outcome of its next project proposal, let us examine the history shown in Figure 1. Here projects have two attributes, x_1 and x_2, the accepted proposals being represented by circles and the rejected proposals by crosses. Top management's utility function $U(x)$ appears in part graphically as a set of iso-utility curves with associated utilities $U_1 < U_2 < U_3 \; . \; . \; . < U_n$. It is apparent from this history, assuming that top management's decisions have been consistent, that $U_k < U^* < U_m$. One may define an acceptance region $A = \{x/U(x) \geq U^*\}$. If U is increasing in x

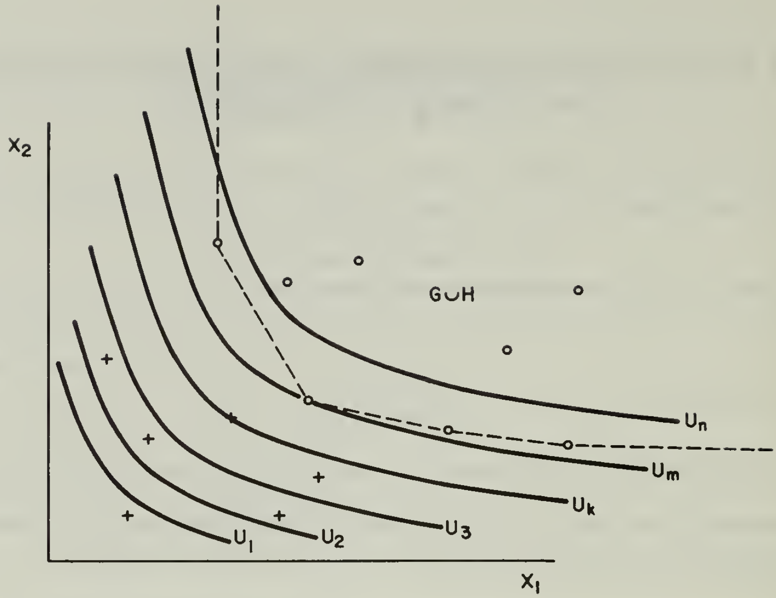

FIGURE 1

and has the property of nonincreasing (absolute) marginal rates of substitution,† then A will be a convex set. Consider every convex combination G of previously accepted proposals x^j; $G=\left\{x \mid x=\sum_j \lambda_j x^j;\ \lambda_j \epsilon [0,1], \sum_j \lambda_j = 1\right\}$. Since $x^j \epsilon A$ and A is convex, then G must be included in A. In addition, since utility is increasing in each of its arguments, the set of points H which dominate any accepted proposal ($H = \{x \mid x \geqslant x^j$, any $x^j\}$) is included in A. Thus a sufficient condition for any new proposal x^0 to be accepted by top management, i.e., for $x^0 \epsilon A$, is that $x^0 \epsilon G \cup H$, where $G \cup H \subseteq A$, or that x^0 be representable as some convex combination of previously accepted proposal vectors or score at least as high on every attribute as some previously accepted proposal. It is obvious that the validity of this argument extends to projects with any number of attributes (dimensions). Moreover, it is evident that A will be convex for any cardinal utility function which is quasi-concave.††

Consider next the history of proposal outcomes shown in Figure 2. Here an apparent inconsistency exists as the previously rejected proposal labeled r is included in G and hence in A. Such an inconsistency can be resolved by one or both of the following explanations:

a. project vectors are incomplete; some relevant piece of information has been left out. All convex combinations of proposals G in some higher dimensional space would exclude r.

b. top management is imperfect in making decisions, such behavior being, perhaps, stochastic in in nature. Errors may be made due to insufficient knowledge of the utility function or due to incorrect perception of proposal attribute magnitudes.

†Marginal rate of substitution of attribute 1 for attribute 2 is defined as $-\partial x_1/\partial x_2 \mid U =$ constant. Thus we require $\partial^2 x_1/\partial x_2^2 \geqslant 0$ for any pair of attributes x_1, x_2. These are fairly general conditions which appear frequently in the economics literature.

††This follows directly from a definition of quasi-concavity (Zangwill [2], p. 34).

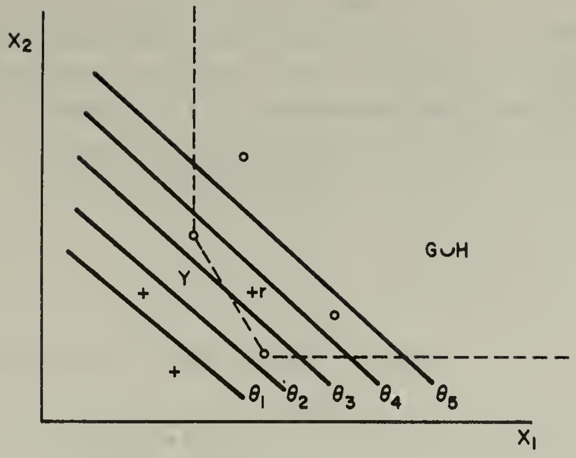

FIGURE 2

The solution to the dilemma in the first instance is straightforward: more complete description of proposals is necessary. In the second case a more detailed model of top management's decision process may be useful. One such model is discussed in the remainder of this paper.

BAYESIAN APPROACH TO PREDICTION

A critical element in the acceptance decision for any new proposal y is the extent of the region A. Hence the division manager seeks an estimate of the position of A's boundary, $B=\{x|U(x)=U^*\}$, given the observed history h of previous proposal outcomes. The following conditions permit us to determine a probability distribution for $B, f_{\tilde{B}/h}(B)$:

(1) B is known to be of form $B(\theta)$, where θ is some unknown parameter (θ may be a vector). In addition, the boundary B for each value of θ is unique; that is, no two different values of θ give the same B.

(2) The divisional manager can provide subjective prior probability estimates $f_{\tilde{\theta}}(\theta)$, the probability that $\tilde{\theta}=\theta$.

(3) Errors are of two types: acceptance of a proposal when $U(x)<U^*$ and rejection of a proposal when $U(x)\geqslant U^*$, and the probabilities of all such errors depend only upon x and θ (or equivalently upon x and B). Error-making behavior can be summarized by the relationship of probability of acceptance P to x and θ, $P=P(x,\theta)$. It should be noted that these are top management errors, not errors on the part of the divisional manager. We assume that $P(x,\theta)$ is known; however, in the most general case Bayesian estimators of the parameters of this function could also be derived. A likelihood function L representing the probability of any history h given θ can then be constructed.

In the Bayesian analysis of the problem, the subjective distribution $f_{\tilde{\theta}}(\theta)$ is combined with the likelihood function $L(h/\theta)$ of the project proposal history from (3) for each of the possible values of θ in such a fashion as to yield a posterior distribution for $\theta, f_{\tilde{\theta}/h}(\theta)$. Since from (1) $B=B(\theta)$, then $f_{\tilde{B}/h}(B)=f_{\tilde{\theta}/h}(\theta)$ for $B=B(\theta)$. The probability of top management accepting any new proposal y is thus $\sum_{\theta} P(y,\theta)\, f_{\tilde{\theta}/h}(\theta)$ if the problem is formulated with discrete probability (mass) function, or $\int_{\theta} P(y,\theta)\, f_{\tilde{\theta}/h}(\theta)d\theta$ if formulated with continuous (density) function.

A SIMPLE EXAMPLE

Let us examine again the history of investment project proposal outcomes represented in Figure 2. Utility is given by $U=\phi(x_1+x_2)$ where ϕ is some strictly increasing function. In this case $B(\theta) = \{x/x_1+x_2=\theta\}$. We wish to determine the probability of a new proposal (shown as y) being accepted. The division manager provides the following judgmental probabilities:

$$f_{\tilde{\theta}}(\theta_j) = \begin{cases} 0.1, & j=1 \\ 0.1, & j=2 \\ 0.5, & j=3 \\ 0.2, & j=4 \\ 0.1, & j=5 \end{cases}$$

Top management's probability of accepting a proposal with utility less than U^* or of rejecting a proposal with utility greater than U^* are both equal to 0.1. This process is illustrated in Figure 3.

Likelihood values representing the probability of this particular proposal history given that $\tilde{\theta}=\theta_j$ are given in this case by $L(h/\theta_j)=0.1^a\,0.1^b\,0.9^c$, where a is the number of apparently erroneous acceptances, b is the number of apparently erroneous rejections, and c is the number of apparently correct decisions. Calculated likelihood values are given below.

$$L(h/\theta_j) = \begin{cases} 0.9^5\,0.1^2 = 0.0059, & j=1 \\ 0.9^6\,0.1 \;= 0.0531, & j=2 \\ 0.9^5\,0.1^2 = 0.0059, & j=3 \\ 0.9^5\,0.1^2 = 0.0059, & j=4 \\ 0.9^4\,0.1^3 = 0.0007, & j=5 \end{cases}$$

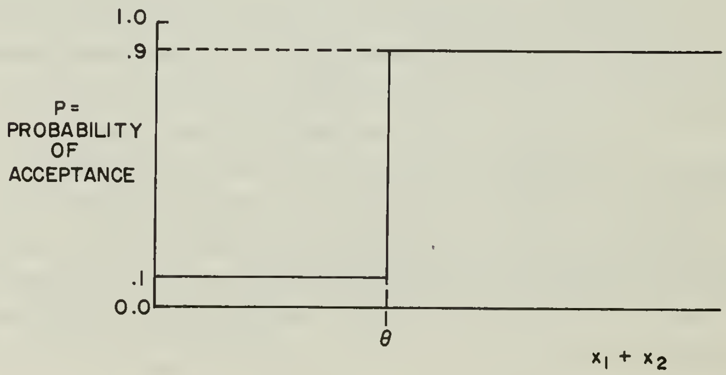

FIGURE 3

The posterior distribution $f_{\tilde{\theta}/h}(\theta)$ is given by Bayes' formula

$$f_{\tilde{\theta}/h}(\theta)=\frac{f_{\tilde{\theta}}(\theta_j)\ L(h/\theta_j)}{\sum_j f_{\tilde{\theta}}(\theta_j)\ L(h/\theta_j)}=\begin{cases}0.059, & j=1\\ 0.526, & j=2\\ 0.292, & j=3\\ 0.117, & j=4\\ 0.006, & j=5\end{cases}$$

and the probability of new proposal y being accepted is

$$0.9f_{\tilde{\theta}/h}(\theta_1)+0.9f_{\tilde{\theta}/h}(\theta_2)+0.1f_{\tilde{\theta}/h}(\theta_3)+0.1f_{\tilde{\theta}/h}(\theta_4)+0.1f_{\tilde{\theta}/h}(\theta_5)=0.57$$

(approximately).

SCALING OF PROPOSALS

From the standpoint of both computational effort and of obtaining reliable judgmental probabilities it is desirable to minimize the number of parameters which define the acceptance boundary B. In the previous example a symmetrical boundary relation $\theta=x_1+x_2$ was employed with the single parameter θ to be estimated. Simple relations requiring a single parameter and providing symmetry about the axis $x_i=x_k \forall i, k$ include $\theta=\sum_i x_i$ and $\theta=\prod_i x_i$. Such a standardized boundary relation may be employed if an appropriate scaling of proposal vectors is made before use.

With x^k representing the scaled proposal corresponding to raw proposal z^k, one possible method of scaling would be to let $x_i^k=z_i^k/\bar{z}_i$, where $\bar{z}_i$ is the arithmetic mean over all known past proposals k for attribute i. This transformation has the very desirable property that element-by-element inequality relationships are preserved; thus, if $z_i^s \geqslant z_i^t$, then $x_i^s \geqslant x_i^t$. In addition, the cluster of scaled scores is symmetric in the sense that the means of all proposals along each attribute dimension are the same (in fact they are each equal to 1). As a practical matter, such information helps provide also a needed frame of reference to the division manager for his development of a meaningful judgmental probability distribution for θ.

SUMMARY

The intent of this discussion has been to suggest a means of prescreening investment project proposals given a limited and apparently inconsistent history and with a relatively simple model of top management behavior in mind. The outcome of the analysis for the division manager is not a "go-no-go" rule for evaluating future proposals, but rather an estimate, assuming that top management's behavior remains unchanged, of the probability of acceptance of the next proposal.

ACKNOWLEDGMENTS

The author is indebted to his colleague Professor William R. Sherrard and to Professor Michael H. Wagner of the New York University Graduate School of Business for helpful comments on an earlier

version of this paper. The author thanks also an (anonymous) referee for suggesting several improvements in the original manuscript.

REFERENCES

[1] Lusk, Edward, "Discriminant Analysis as Applied to the Resource Allocation Decision," The Accounting Review (July 1972).

[2] Zangwill, W. I., *Nonlinear Programming: a Unified Approach* (Prentice-Hall, 1969).

MAXIMIZATION OF LONG-RUN AVERAGE RATE-OF-RETURN BY STOCHASTIC APPROXIMATION

Robert A. Agnew

Montgomery Ward
Chicago, Illinois

ABSTRACT

Suppose that an individual has a surplus stock of wealth and a fixed set of risky investment opportunities over a sequence of time periods. Assuming the criterion of maximal long-run average rate-of-return, the individual may select portfolios sequentially via a modified stochastic approximation procedure. This approach yields optimal asymptotic investment results under minimal assumptions.

1. INTRODUCTION

Suppose that an individual has a stock of wealth which is not required for short-run consumption expenditures and a fixed set of risky investment opportunities over an indefinite horizon of discrete time periods. Assuming time homogeneity and period-wise independence (or near independence) of investment results, Breiman [6], [7] has demonstrated that a portfolio (i.e., an allocation of wealth to the various investment opportunities) which maximizes the expected logarithm of period-wise growth is "optimal" in the sense that it maximizes the long-run average rate-of-return (or rate of growth) as the time horizon tends to infinity. (Actually, the treatment in [6] is somewhat more general.) Moreover, Breiman's optimal portfolio is a "fixed proportions" portfolio in that the proportions of wealth allocated to the various investment opportunities do not vary with the magnitude of wealth.

Now, if the individual accepts Breiman's optimality criterion, he will want to obtain the optimal proportions and employ them at every future period. However, in any realistic situation, the individual can only estimate the optimal proportions initially, and he will want to improve his estimates (and hence his investment results) as he goes along. A good sequential approximation scheme should yield the optimal proportions and should attain the maximal average rate-of-return in the limit as the time horizon tends to infinity. (It appears that even the best conceivable statistical procedures will have sufficient asymptotic variation to render them "inadmissible" in the sense of Breiman [6]; i.e., the individual would be infinitely better off if he knew the optimal proportions. Of course, our purpose is to treat the case where these optimal proportions are unknown.)

As a suitable approximation scheme, we suggest a modified (generally multivariate) Kiefer-Wolfowitz stochastic approximation procedure [24], [10] to [15], [32], [33], [37], [38]. This approach requires minimal assumptions and calculations, and it yields the desired results in the limit. After some preliminaries in the next section, the approximation procedure is presented in section 3 and its convergence properties are discussed. A simulated numerical example is presented in section 4, and section 5 discusses the addition of withdrawals for consumption.

2. PRELIMINARIES

Let $G_1, \ldots, G_k$ be nonnegative random variables representing the "growth factors" associated with $k \geq 1$ risky investment opportunities over a given time period; i.e., a dollar invested in opportunity j yields G_j dollars at the end of the period. We assume that $E(G_j^2) < \infty$ for all $j = 1, \ldots, k$.

Let x_j be the proportion of wealth invested in opportunity j. We shall require that $\sum_{j=1}^{k} x_j \leq \gamma$ for some $\gamma \epsilon (0, 1)$; $1 - \sum_{j=1}^{k} x_j \geq 1 - \gamma > 0$ then represents the proportion of wealth held in a "riskless asset" assumed to have growth factor $g \geq 1$ with probability one. Let

$$S = \left\{ x = (x_1, \ldots, x_k) : \sum_{j=1}^{k} x_j \leq \gamma; \qquad x_1, \ldots, x_k \geq 0 \right\} \tag{1}$$

represent the set of feasible proportional allocations or "portfolios." Given $x \epsilon S$, let

$$G(x) = g\left(1 - \sum_{j=1}^{k} x_j\right) + \sum_{j=1}^{k} G_j x_j = g + \sum_{j=1}^{k} (G_j - g) x_j \tag{2}$$

represent the portfolio growth factor over the given period; i.e., a dollar invested in portfolio $x \epsilon S$ yields $G(x)$ dollars at the end of the period. According to Breiman's criterion, the objective is to maximize

$$f(x) = E(\log (G(x))) \tag{3}$$

over S. This expectation exists because $G(x) \geq g(1-\gamma) > 0$ and $E(G(x)) < \infty$. In fact, it is easy to see that $E(|\log (G(x))|^p)$ is uniformly bounded on S for any integer $p \geq 1$.

By the strict concavity of the log function, f is concave on S. Moreover, it is strictly concave if the $G_j - g$ are linearly independent (i.e., $P\left\{\sum_{j=1}^{k} (G_j - g)\xi_j = 0\right\} < 1$ whenever $\xi = (\xi_1, \ldots, \xi_k) \neq 0$), and this latter condition holds if the covariance matrix of the G_j is positive definite. We shall assume strict concavity. Furthermore, we shall assume that f attains its unique maximum at an interior point $\theta \epsilon S$; i.e., none of our investment opportunities, including the riskless asset, is completely inferior to any other, and the optimal portfolio is "diversified." Since the riskless portfolio is feasible, it follows that $f(\theta) > \log (g) \geq 0$.

Under our previous assumptions, the first and second order partial derivatives of f exist and are uniformly bounded in absolute value on S. Indeed, we have

$$\partial f / \partial x_j (x) = E((G_j - g)/G(x)) \tag{4}$$

so the optimality condition $\partial f / \partial x_j (\theta) = 0$ implies

$$E(G_j / G(\theta)) = E(g / G(\theta)) = 1 \tag{5}$$

which indicates that the optimal portfolio "balances" the various investment opportunities in an intuitive way. The second order derivatives

$$\partial^2 f/\partial x_i \partial x_j(x) = -E((G_i - g)(G_j - g)/G(x)^2) \tag{6}$$

need not be continuous, but continuity would follow from the existence of higher moments for the G_j. In general, the existence of pth order moments for the G_j implies the existence and boundedness of pth order derivatives for f. If, in particular, the G_j are bounded random variables, derivatives of all orders exist on S.

3. THE APPROXIMATION PROCEDURE

Let $W_0 > 0$ be our initial wealth, and let $X_1 = (X_{1,1}, \ldots, X_{1,k})$ be our initial estimate of θ. We assume that $c < \min_j X_{1,j}$ and $\sum_{j=1}^{k} X_{1,j} < \gamma - c$, where $0 < c < \gamma/(k+1)$. Define positive sequences $\{a_n\}$ and $\{c_n\}$ by $a_n = an^{-\alpha}$ and $c_n = cn^{-\beta}$, where $0 < \beta < 1/2$ and $\max(1/2 + \beta, 1 - 2\beta) < \alpha < 1$. Let e_j denote the jth unit vector of dimension k, and let $h(u) = 0$ for $u \leq 0$, $h(u) = u$ for $u > 0$.

Our approximation procedure will evolve over a sequence of "stages," each stage encompassing $2k$ consecutive time periods. Our estimate of θ at the nth stage will be denoted by $X_n = (X_{n,1}, \ldots, X_{n,k})$, and our wealth after the nth stage will be denoted by W_n. During the nth stage, we implement the $2k$ portfolios $X_n \pm c_n e_j$ $(j = 1, \ldots, k)$, and we realize the corresponding growth factors

$$G^{\pm}_{n,j} \overset{d}{=} G(X_n \pm c_n e_j), \tag{7}$$

where $\overset{d}{=}$ means "identically distributed," and $G(x)$ denotes the random function defined in section 2. It follows then that

$$E(|\log(G^{\pm}_{n,j})|^2) \leq C < \infty \tag{8}$$

and assuming period-wise independence (or near independence) of component growth factors

$$E(\log(G^{\pm}_{n,j})|X_n, \ldots, X_1) = f(X_n \pm c_n e_j) \tag{9}$$

so the Kiefer-Wolfowitz scheme would ordinarily put

$$X_{n+1} = X_n + a_n Y_n, \tag{10}$$

where $Y_n = (Y_{n,1}, \ldots, Y_{n,k})$, and

$$Y_{n,j} = (2c_n)^{-1}(\log(G^{+}_{n,j}) - \log(G^{-}_{n,j})) \tag{11}$$

is an estimate of $\partial f/\partial x_j(X_n)$. However, we must modify the basic recursion relation (10) in order to ensure feasibility. Our proposed modification is in no sense unique, but it is rather simple.

We put

$$J_n = \{j : Y_{n,j} > 0\}, \tag{12}$$

$$A_{n,j} = Y_{n,j} \Big/ \sum_{i \in J_n} Y_{n,i} \qquad (j \in J_n), \tag{13}$$

$$B_{n,j} = \max\,(c_n, X_{n,j} + a_n Y_{n,j}), \tag{14}$$

$$X_{n+1,j} = B_{n,j} > c_{n+1} \qquad (j \notin J_n), \tag{15}$$

$$X_{n+1,j} = B_{n,j} - A_{n,j} h \left(\sum_{i=1}^{k} B_{n,i} - \gamma + c_n \right) \qquad (j \in J_n), \tag{16}$$

and it is not difficult to establish

$$X_{n+1,j} > X_{n,j} \qquad (j \in J_n), \tag{17}$$

$$\sum_{j=1}^{k} X_{n+1,j} \leq \gamma - c_n < \gamma - c_{n+1}, \tag{18}$$

from which it follows that all implemented portfolios $X_n \pm c_n e_j$ are interior points of the feasible set S.

Since θ is an interior point of S by assumption, the asymptotic behavior of the approximation procedure will be unaffected by the feasibility requirements. Under our assumptions (cf. Theorem 4.4 in [33]), $X_n \to \theta$ with probability one as $n \to \infty$. Moreover,

$$\log(W_n) - \log(W_{n-1}) = \sum_{j=1}^{k} (\log(G_{n,j}^{+}) + \log(G_{n,j}^{-})) \tag{19}$$

so that $n^{-1}(\log(W_n) - \log(W_0)) \to 2kf(\theta)$ or $W_n^{1/n} \to \exp(2kf(\theta))$ with probability one by the Strong Law for Martingales [18]. Hence, the optimal proportions and the maximal average rate-of-return are attained almost surely in the limit as the time horizon tends to infinity.

Now, a, c, α, β, γ, and the vector X_1 are parameters to be stipulated by the individual. Obviously, X_1 should be the individual's best initial estimate of θ, and γ should be chosen sufficiently close to one so that θ is certain to be an interior point of the feasible set S. Furthermore, it is clear that a should be sufficiently small relative to c so that the procedure will avoid the boundary at the first iteration. (Indeed, a can be stipulated after the first stage results are known.)

Under slightly stronger assumptions, we can investigate the effect of various choices of a, c, α, and β on the asymptotic behavior of our approximation procedure. Suppose, for instance, that our basic growth factors in section 2 possess finite third moments so that f has bounded third derivatives on S. In addition, suppose that the minimum eigenvalue of $-H(x)$ is bounded away from zero on S, where

(20) $$H(x) = [\partial^2 f/\partial x_i \partial x_j (x)]_{i,j=1}^{k}$$

is the Hessian of f. Under these assumptions, $0 < \beta < 1/6$ and $\alpha = 6\beta$ yield

(21) $$E(||X_n - \theta||^2) = 0(n^{-4\beta})$$

(cf. Theorem 4.6 in [33]) so the procedure converges in the mean square even if the various conditions on α and β for probability one convergence are violated. Indeed, for β very close to zero and $\alpha = 6\beta$, our procedure is akin to a perpetual control process.

Assuming further that f has continuous third derivatives and that $\alpha = 6\beta$ with $1/8 < \beta < 1/6$, we have $X_n \to \theta$ with probability one, and in addition $n^{2\beta}(X_n - \theta)$ is asymptotically normal (cf. Theorem 4.7 in [33]) with mean vector

(22) $$\mu = c^2 m$$

and covariance matrix

(23) $$V = (a/c^2)\sigma^2 M,$$

where

(24) $$\sigma^2 = \text{var}(\log(G(\theta)))$$

and m is a vector, M a matrix, dependent only on the behavior of f in the immediate vicinity of θ. It follows that asymptotically

(25) $$E(||X_n - \theta||^2) \simeq n^{-4\beta}(b_1 c^4 + b_2(a/c^2)),$$

where b_1 and b_2 are positive constants which are independent of the choice of parameters and which are in some sense inversely proportional to the degree of curvature in f at θ. From this result, it is clear that asymptotic speed and precision are increased as β is increased, as c is decreased, and as a is decreased relative to c. On the other hand, if a is chosen too small, the procedure will barely move at all in the short run, and this may be undesirable if X_1 is not a particularly good estimate of θ. In general, the individual must stipulate the parameters a, d, and β subjectively in order to balance his short and long term expectations.

4. SIMULATED NUMERICAL EXAMPLE

Consider a simplified environment where cash is the riskless asset with $g = 1$ and the single risky investment opportunity is a simple "double or nothing" favorable gamble (cf. [19], [35], [36]) with constant probability success $p = 0.55$. Then,

(26) $$f(x) = p \log(1 + x) + q \log(1 - x)$$

with $q=1-p=0.45$. Furthermore,

$$\theta=p-q=0.10 \tag{27}$$

$$\exp(f(\theta))=2p^pq^q \simeq 1.005 \tag{28}$$

$$f''(\theta)=-1/4pq \simeq -1.010 \tag{29}$$

$$f'''(\theta)=-\theta/4(pq)^2 \simeq -0.408 \tag{30}$$

$$\sigma^2=pq(\log(p/q))^2 \simeq 0.010 \tag{31}$$

and the constants in (25) are

$$b_1=(f'''(\theta)/6f''(\theta))^2 \simeq 0.0045 \tag{32}$$

$$b_2=\sigma^2/4|f''(\theta)| \simeq 0.0025. \tag{33}$$

Now suppose that our individual has 50 independent gambling opportunities per year so that the optimum annualized long-run rate-of-return is about 28.4 percent. Since $k=1$, 50 trials correspond to 25 stages per year in our approximation procedure, and we may simplify the notation of section 3 to

$$G_n^{\pm}=1+(2U_n^{\pm}-1)(X_n \pm c_n) \tag{34}$$

$$Y_n=(2c_n)^{-1}(\log(G_n^+)-\log(G_n^-)) \tag{35}$$

$$X_{n+1}=\max(c_n, \min(\gamma-c_n, X_n+a_nY_n)) \tag{36}$$

$$W_n=\exp\left(\sum_{i=1}^{n}(\log(G_i^+)+\log(G_i^-))\right) \tag{37}$$

where $X_n \pm c_n \in (0, \gamma)$ are simply the proportions of wealth gambled at the two trials in stage n, $U_n^{\pm}$ are the corresponding zero/one (lose/win) indicator random variables, and $W_0 \equiv 1$ for simplicity.

In order to better illustrate how our approximation procedure works, we have simulated 10 1-year runs with $X_1=0.08$, $c=0.05$, $a=0.001$ and 0.01, $\gamma=0.20$, $\beta=0.15$, and $\alpha=6\beta=0.90$. The random numbers employed were taken from the first 10 columns on p. 480 of [4]; i.e., each column of 50 random numbers was used to generate a 1-year simulation run. The net results are listed in Table 1, whereas Run 2 is detailed step-by-step in Table 2. Note the difference in short-run variation between the cases $a=0.001$ and $a=0.01$.

5. CONCLUDING REMARKS

Suppose that our individual wishes to make withdrawals from his stock of wealth for purposes of consumption. Joint optimization of consumption-investment decisions over time has generally been

TABLE 1. *Net Simulation Results*

Run	$a=0.001$		$a=0.01$	
	X_{26}	W_{25}	X_{26}	W_{25}
1	0.0825	2.997	0.1044	3.785
2	0.0816	1.364	0.0958	1.334
3	0.0789	1.281	0.0679	1.229
4	0.0817	1.288	0.0966	1.220
5	0.0776	2.208	0.0602	1.698
6	0.0788	3.612	0.0694	2.779
7	0.0764	0.409	0.0486	0.480
8	0.0817	1.124	0.0972	1.047
9	0.0820	1.715	0.0993	1.691
10	0.0818	1.693	0.0970	1.735
Average	0.0803	1.769	0.0836	1.700

TABLE 2. *Simulation Run 2*

n	U_n^+	U_n^-	$a=0.001$		$a=0.01$	
			X_n	W_n	X_n	W_n
1	1	1	0.0800	1.164	0.0800	1.164
2	1	0	0.0809	1.264	0.0893	1.262
3	1	1	0.0818	1.477	0.0994	1.523
4	0	0	0.0822	1.241	0.1028	1.223
5	1	1	0.0819	1.451	0.0996	1.477
6	0	1	0.0821	1.333	0.1018	1.351
7	0	0	0.0817	1.122	0.0962	1.102
8	0	1	0.0815	1.034	0.0943	1.013
9	1	1	0.0811	1.207	0.0902	1.203
10	1	0	0.0812	1.286	0.0914	1.279
11	0	0	0.0815	1.083	0.0946	1.047
12	1	1	0.0814	1.266	0.0933	1.250
13	1	0	0.0815	1.345	0.0943	1.326
14	1	0	0.0817	1.428	0.0970	1.404
15	1	1	0.0819	1.670	0.0996	1.696
16	1	0	0.0820	1.771	0.1004	1.793
17	0	0	0.0822	1.490	0.1028	1.441
18	0	1	0.0821	1.384	0.1019	1.334
19	0	1	0.0819	1.288	0.0995	1.237
20	0	0	0.0817	1.084	0.0972	1.007
21	1	0	0.0817	1.147	0.0965	1.062
22	1	1	0.0818	1.341	0.0984	1.280
23	1	1	0.0819	1.569	0.0990	1.545
24	0	1	0.0819	1.462	0.0995	1.435
25	0	1	0.0818	1.364	0.0976	1.334
26			0.0816		0.0958	

addressed in the literature via two criteria: maximum discounted expected utility of long-run consumption (or a suitable combination of consumption and wealth) [17], [21], [22], [25], [27], [28], [30] and minimum probability of ruin in the context of rigid consumption requirements [19], [35], [36].

Both criteria yield rather complicated dynamic programming type functional equations to be solved for the optimum consumption-investment policies. (It should be mentioned, however, that a variety of intriguing inferential results have been obtained via these characterizations.) We shall ignore the problem of joint optimization and merely indicate an intuitive means of adding a flexible consumption policy to the approximation procedure developed in section 3.

Let ρ_n be a nonnegative sequence converging to zero (e.g., $\rho_n \equiv 0$). Referring to section 3, we put

$$Z_n = \sum_{j=1}^{k} (\log (G_{n,j}^{+}) + \log (G_{n,j}^{-})) \tag{38}$$

$$V_n = n^{-1} \sum_{i=1}^{n} Z_i = n^{-1} Z_n + (1 - n^{-1}) V_{n-1} \tag{39}$$

$(V_0 = 0)$ so that $V_n \to 2kf(\theta)$ with probability one. Putting

$$T_n = \max (\rho_n, V_n), \tag{40}$$

we also have $T_n \to 2kf(\theta)$ with probability one. Now let C_n be the withdrawal for consumption at the end of stage n and let W_n be the postwithdrawal wealth so that

$$W_n + C_n = W_{n-1} \exp (Z_n). \tag{41}$$

Our policy is

$$W_n = (W_n + C_n) \exp (-\delta T_n) \tag{42}$$

$$C_n = (W_n + C_n)(1 - \exp (-\delta T_n)), \tag{43}$$

where $\delta \epsilon (0, 1)$ is a subjective parameter. (The procedure in section 3 corresponds to $\delta = 0$.)

We have then that $W_n^{1/n}$ and also $C_n^{1/n}$ converge to $\exp ((1 - \delta) 2kf(\theta))$ with probability one so that δ provides a flexible mechanism for trading off short and long term consumption. The long term is weighted heavily when δ is near zero, whereas the short term is weighted heavily when δ is near one. (Note that our stage-wise withdrawal policy could easily be made period-wise.)

Given an extensive set of risky investment opportunities, it may be necessary to aggregate in order to render the approximation problem manageable. For instance, one can imagine a one-dimensional trade-off between a riskless asset and a suitably defined "market portfolio" [23].

Regarding speed of convergence, we note that asymptotic speed can be improved via Fabian's generalized Kiefer-Wolfowitz procedures at the expense of increased analytical complexity [13], [14], [15], [33]. Nevertheless, any statistical procedure will converge rather slowly in comparison to deterministic procedures. On the other hand, deterministic procedures do not apply directly to problems where randomness is a fundamental ingredient. Any rigorous statistical procedure will proceed rather cautiously in order to properly screen valid information about the underlying objective function from pure random noise [38].

Samuelson [31] has questioned the long-run average rate-of-return criterion on the basis of alternative, finite-stage utility criteria. He notes that not all individuals need be limiting logarithmic utility maximizers. While this is undoubtedly true, logarithmic utility is distinguished by the fact that it does possess some external justification. On the other hand, one can easily tailor our approximation procedure to a modified criterion such as

$$f(x)=E(G(x)^{\eta}) \tag{44}$$

$(0<\eta<1)$ which would maximize $E((W_n/W_{n-1})^{\eta})$ in the limit. In any event, the infinite stage setup is obviously convenient when the ultimate number of stages is undetermined.

BIBLIOGRAPHY

[1] Agnew, R. A., "Counter-Examples to an Assertion Concerning the Normal Distribution and a New Stochastic Price Fluctuation Model," Review of Economic Studies *38*, 381–383 (1971).

[2] Agnew, R. A., "Sequential Bid Selection by Stochastic Approximation," Nav. Res. Log. Quart. *19*, 137–143 (1972).

[3] Agnew, R. A., "Inequalities with Application in Economic Risk Analysis," J. Applied Probability *9*, 441–444 (1972).

[4] Beyer, W. H., Editor, *CRC Handbook of Tables for Probability and Statistics* (The Chemical Rubber Co., Cleveland, O., 2nd Ed. 1968).

[5] Borch, K. H., *The Economics of Uncertainty* (Princeton University Press, Princeton, N.J., 1968).

[6] Breiman, L., "Investment Policies for Expanding Businesses Optimal in a Long-Run Sense," Nav. Res. Log. Quart. *7*, 647–651 (1960).

[7] Breiman, L., "Optimal Gambling Systems for Favorable Games," Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability *1*, 65–78 (1961).

[8] Comer, J. P., Jr., "Some Stochastic Approximation Procedures for use in Process Control," Annals of Mathematical Statistics *35*, 1136–1146 (1964).

[9] Comer, J. P., Jr., "Application of Stochastic Approximation to Process Control," J. Roy. Statist. Soc. *B27*, 321–331 (1965).

[10] Dupač, V., "On the Kiefer-Wolfowitz Approximation Method," Casopis Pest. Mat. *82*, 47–75 (1957). English translation in *Selected Translations in Mathematical Statistics and Probability*, Am. Math. Soc., Vol. IV (1963).

[11] Dvoretsky, A., "On Stochastic Approximation," Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability *1*, 39–55 (1956).

[12] Fabian, V., "Stochastic Approximation of Constrained Minima," Transactions of the Fourth Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, 277–290 (1967).

[13] Fabian, V., "Stochastic Approximation of Minima with Improved Asymptotic Speed," Annals of Mathematical Statistics *38*, 191–200 (1967).

[14] Fabian, V., "On the Choice of Design in Stochastic Approximation Methods," Annals of Mathematical Statistics *39*, 457–465 (1968).

[15] Fabian, V., "On Asymptotic Normality in Stochastic Approximation," Annals of Mathematical Statistics *39*, 1327–1332 (1968).

[16] Fama, E. F., "The Behavior of Stock Market Prices," Journal of Business *38*, 34–105 (1965).

[17] Fama, E. F., "Multiperiod Consumption—Investment Decisions," American Economic Review *60*, 163–174 (1970).

[18] Feller, W., *An Introduction to Probability Theory and its Applications* (John Wiley and Sons, Inc., N.Y., Vol. I, 1968, 3rd Ed.; Vol. II, 2nd Ed., 1971).

[19] Ferguson, T. S., "Betting Systems which Minimize the Probability of Ruin," Journal of the Society for Industrial and Applied Mathematics *13*, 795–818 (1965).

[20] Hakansson, N. H. and Tein-Ching Liu, "Optimal Growth Portfolios When Yields are Serially Correlated," Review of Economics and Statistics *52*, 385–394 (1970).

[21] Hakansson, N. H., "Optimal Investment and Consumption Strategies under Risk for a Class of Utility Functions," Econometrica *38*, 587–607 (1970).

[22] Hakansson, N. H., "Optimal Entrepreneurial Decisions in a Completely Stochastic Environment," Management Science *17*, 427–449 (1971).

[23] Jensen, M. C., "Capital Markets: Theory and Evidence," Bell Journal of Economics and Management Science *3*, 357–398 (1972).

[24] Kiefer, J. and J. Wolfowitz, "Stochastic Estimation of the Maximum of a Regression Function," Annals of Mathematical Statistics *23*, 462–466 (1952).

[25] Levhari, D. and T. N. Srinivasan, "Optimal Savings under Uncertainty," Review of Economic Studies *36*, 153–163 (1969).

[26] Markowitz, H. M., *Portfolio Selection* (John Wiley and Sons, Inc., N.Y., 1959).

[27] Mossin, J., "Optimal Multiperiod Portfolio Policies," Journal of Business *41*, 215–229 (1968).

[28] Phelps, E. S., "The Accumulation of Risky Capital: A Sequential Utility Analysis," Econometrica *30*, 729–743 (1962).

[29] Pickett, J. R., "The Random Walk Theory of Stock Price Behavior," Unpublished M. S. Thesis, Air Force Institute of Technology (1970).

[30] Samuelson, P. A., "Lifetime Portfolio Selection by Dynamic Stochastic Programming," Review of Economics and Statistics *51*, 239–246 (1969).

[31] Samuelson, P. A., "The 'Fallacy' of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling," Proceedings of the National Academy of Sciences *68*, 2493–2496 (1971).

[32] Schmetterer, L., "Stochastic Approximation," Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability *1*, 587–609 (1961).

[33] Schmetterer, L., "Multidimensional Stochastic Approximation," in *Multivariate Analysis II* (Academic Press, N.Y., Ed. by P. R. Krishnaiah, 1969).

[34] Tower, T. R., "Sequential Inventory Control and Optimization through Stochastic Approximation," Unpublished M.S. Thesis, Air Force Institute of Technology (1972).

[35] Truelove, A. J., "A Multistage Stochastic Investment Process," Rand Research Memorandum RM–4025–PR (1964).

[36] Truelove, A. J., "Betting Systems in Favorable Games," Annals of Mathematical Statistics *41*, 551–566 (1970).

[37] Wasan, M. T., *Stochastic Approximation* (Cambridge University Press, 1969).

[38] Wilde, D. J., *Optimum Seeking Methods* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964).

[39] Zangwill, W. I., *Nonlinear Programming: A Unified Approach* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969).

STRUCTURAL MODELS OF AWARD FEE CONTRACTS

Clifford W. Marshall

Polytechnic Institute of New York

ABSTRACT

This paper describes the construction of a mathematical model structure in a particular area of management. In addition it is intended as an illustration of how appropriate levels of mathematics can be utilized in management research when original formulations lack sufficient precision for conducting quantitative analyses. The specific area studied deals with the formulation and analysis of contract types. In particular the award fee type contract is treated. At present mathematical structure models for other contract types have received considerable attention, but award fee types have not been structured in mathematical terms. The paper provides a discussion of model formulation for award fee contracts, develops a detailed example of such a structure, and illustrates that model by numerical examples indicating the application of such models to the formulation and analysis of award fee contracts.

INTRODUCTION

In order to carry out their functional objectives, departments of the Federal Government and other large organizations commonly employ a wide variety of contractual arrangements with their contractors.* That variety is needed to accommodate the many relevant factors involved, such as risk assumption, value of goods or services, payment of fair profits, and motivation of contractors. The vast majority of contract types may be divided into either incentive or nonincentive fee types. Special technical considerations require many varieties of contract types within each of the two major categories, however, intracategory differences are much less than the basic distinction between categories. Nonincentive fee contact types are utilized in dealing with standard situations involving little risk (e.g., Fixed Price Contracts for standard parts) or in extremely uncertain situations (e.g., Cost Plus Fixed Fee Contracts for state-of-the-art research). Such contracts are not closely related to the award fee concept and will not be discussed further.

Incentive fee contracts employ the profit motive to direct the contractor to operate in a manner considered to be most desirable by the buyer. Cost has been the most commonly "incentivised" contractual factor. Cost incentives are designed to motivate the contractor to keep costs down (as opposed to Cost Plus Fixed Fee contracts for example, that place no reward on cost reduction effort). Such incentive contracts have a great deal of appeal for the contractee and also many attractive features from the contractor's point of view (such as providing an opportunity for it to utilize skillful management). Therefore incentive contracts have enjoyed extensive use over the past decade.† This

*There exists an extensive literature in which contract types are discussed. Examples from that literature are: Peck and Scherer [14], Scherer [16], DOD and NASA contracting guides such as [12] and [13], and Marshall [9].

†Incentive contracts have been employed for a long time. A classical example is the Government's contract with the Wright Brothers. Such contracts came into common use in the post World War II era and have been used increasingly since that time.

is particularly true in areas where fixed price contracts are too risky for the contractor and cost plus fixed fee contracts are too conservative (as when necessary technology for a development is already established). Such popularity led to the formulation of so-called multiple incentive contracts in which factors other than cost were incentivised. Schedules were incentivised in much the same way cost had been since they are well defined factors in many contracts. However, performance factors proved far more difficult to incentivise. Multiple incentive contracts that included performance incentives had difficulty from the beginning. That difficulty seems largely due to a combination of problems associated with performance specification and the formulation of appropriate contract structures. An incentive contract motivates the contractor by establishing quantitative rules relating its profit to its achievement of the buyer's objective.† Such quantitative rules are called the contract structure. They are developed by both parties as part of contract negotiation, specifying such things as contractor fee as a function of the incentivised factor in the contract. In the case of simple cost incentive contracts the structure is commonly one or more straight line segments specifying fee as a function of (total) cost. The proper formulation of such models (structures) for single (cost) incentives is difficult. Though the mathematical form of the single incentive structure is simple, specific realization of a contract that will achieve proper motivation has proven difficult in many cases. Most attempts at developing structures for multiple incentive contracts have been based on combinations of several single incentive structures and have assumed independence of incentive interaction effects (though of course tradeoffs are required in the distribution of the total fee pool available).** Even such relatively simple structures have proven difficult to realize in proper motivational terms as discussed in Jones [7]. Many examples exist in which the contractor has operated in undesirable ways within contract specification and achieved increased fee. Recently some large scale effort has been made to help formalize multiple incentive contract structuring with the aid of computerized trade-off analyses and stylized structure forms.†† Such formalism certainly helps though it does not solve the problem of insuring proper contractor motivation. Moreover the formalisms employed use simple structures which more often than not prevent a priori the achievement of ideal motivational structures. Structuring difficulties are compounded by questions of how to define and express performance. Some suthorities insist on a number of specific, measurable performance factors while others maintain that some overall performance index should be incentivised. In view of this situation the theoretical ideals that are present in the incentive fee concept tend to become lost in application. In addition many critics hold that the profit motive is ineffective or inappropriate as a control mechanism when viewed in the incentive contract context. Moreover, statistical studies are showing real difficulties with the proper formulation of incentive contracts (as indicated by the negotiation of high target costs to provide leverage for cost incentive fee).***

†A referee has pointed out that it is useful to realize that in actual situations a contractor may be motivated by larger considerations than optimizing its value for a single contract. This broad objective view is difficult to formulate or analyze, but even so, detailed suboptimal studies of the type presented here may contribute to general understanding.

**Examples of such structuring techniques are STOIC [6] and Department of the Air Force [2]. More complex treatments of multiple incentive structuring are given in Condon [1] and Marshall [10]

††In April 1967 the Assistant Security of Defense/Installations and Logistics deisgnated the U.S. Air Force as Executive Agency to provide a computerized incentive structuring service to all DOD. That service, designated Program Office for Evaluation and Structuring of Multiple Incentive Contracts (POESMIC) is currently operational.

***Studies addressing problems with multiple incentive contracts may be found in Fisher [5] and a Logistics Management Institute report LMI [8].

The problems with incentive fee contracts are not so great as to abandon the use of such contracts. Indeed they work very well in certain situations, particularly when only cost (and possibly schedule) is incentivised. However when there is a desire to motivate the contractor in the area of performance, alternate contract types seem highly desirable. Typical of informed opinion in this connection are Cravens' remarks "I believe that we can change some of our procedures in order to help—with some—problems. Certainly, some of our incentives have been too complex. Sophisticated concepts do not have to be complex."* This situation has led to the creation of so-called award fee contracts which are very different from standard incentive fee types as discussed for example by Cravens [3], and by Cravens and Rule [4].

The National Aeronautics and Space Administration (NASA) guide on award fee contracts [13] states that an award fee contract "provides that the contractor's variable fee will be determined subjectively by designated, high-level, Government personnel on the basis of periodic, after-the-fact evaluations of the contractor's performance."† It further states that "The intent of the (evaluation) criteria should be as clear as possible so as to provide clear goals against which the contractor works," and that "Subjective evaluation will include objective measurements whenever such standards are appropriate and can be defined." NASA has been a leading user of award fee type contracts over the past several years. The Department of Defense and particularly the U.S. Navy has also made use of award fee contracts. Such usage has been primarily for service contracts, though in some cases goods have also been included within an award fee contract.

The award fee is seldom (if ever) used in pure form. It is most often associated with an incentive or fixed fee to form combined contract structures such as Cost Plus Award Fee or Fixed Price Award Fee contracts. In fact the NASA guide also states that "Subjective evaluation (Cost Plus Award Fee) may be combined with formula type (Cost Plus Incentive Fee) incentives." Thus one sees that the award fee concept forms a part of the overall spectrum of contract types.

Formulation of structural models for award fee contracts will be presented here in terms of basic aspects required of such structures, subsequently to be described. Some mathematical models will be given in detail with numerical examples illustrating analysis of such models.

AWARD FEE CONTRACT STRUCTURES

The award fee concept can be used in a pure form or be combined with an incentive fee. It is usual to specify some fixed fee value and to provide the award fee as an additional increment to that fee. Such a division of fee definition is necessary to reduce contractor risk assumption to a level it will accept in view of the inherent risk introduced by the subjective character of award fees.

Models for award fee contracts may be constructed in which the award fee is either binary or graded. Both kinds of award fees are used in contracting. A binary award fee is a stipulated amount that will or will not be paid on the basis of contractor performance. Graded award fees have a range of possible values and the amount selected is based on contractor performance. Only one amount is paid in fee, but the value of that amount is represented by a range of possibilities. In the binary case the random elements is the granting of an award of specified amount while in the graded case the random element is the amount of the award. The award fee is not stipulated in advance and therefore

*Remarks from a talk by James B. Cravens, NASA Special Assistant to Director of Procurement, made at Cape Kennedy, June 6, 1968.

†In fact there is no need to make such evaluations "periodically"; they can be made at any time or on milestone events.

has a strong subjective character. It depends on many factors that are difficult to assess at the time a contract is negotiated. Therefore a possible model of the award fee outcomes can be based on a random fee formulation. The probabilistic nature of that random fee will reflect the unsure nature of anticipated actual fee.

This paper is primarily concerned with graded award fees. Consideration of binary fees is included here to round out the general description of the award fee concept and illustrate the possibilities for structuring award fee contracts.

Awards may be single or multiple. In an award fee contract several award events may be stipulated or only one event may be considered for an award fee. Multiple award fees may be based on calendar award periods or on milestone events without regard to when those events occur.* How the awards are paid over time has no effect on the award fee structuring and will not be discussed further. Multiple award fees may be independent of each other or they may be dependent. The independent case leads to the most simple structures, but may not yield a satisfactory degree of motivation to achieve the contractee's desires or may yield a structure that is unacceptable to the contractor (e.g., it may be too conservative for a highly entrepreneural contractor). Dependent structures provide a wide range of situations in which the two parties can search for negotiable contracts. In principle any kind of structure might be used but in practice it is desriable to have relatively simple structures that can be analyzed in terms closely related to well defined concepts. The mathematical structure for multiple award fee contracts that follows will be specialized to form such dependent structures. Discussion of independent structures will be given and compared with the dependent results.

Let n denote the number of award fees to be included in a contract (the case of a single award fee is a simple special case of the model being defined).† In our model formulation the award fee structure is represented as a stochastic process $\{x_i\}$, $i = 1, \ldots, n$ with joint distribution function

$$F_{X_1, \ldots, X_n}(x_1, \ldots, x_n).$$

The random variable X_i has distribution function

$$F_{X_i}(x_i) = F_{X_1, \ldots, X_n}(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty).$$

Analysis of the award fee structure studies the random variable H_n whose values represent total award fee. In general, a full description of H_n is not possible and it is discussed in terms of its expected value $A_n = E[H_n]$ and variance $S_n = \text{Var}[H_n]$. Thus it is taken as a desirable feature of any contract structure to provide reasonable expressions for A_n and S_n. Such quantities may be related to specific concepts by negotiators and used as numerical parameters for comparing contracts. Any structure that is so complex that these quantities cannot be directly evaluated is not a useful model.

Binary award fee strauctures correspond to letting $X_i = 1$ when the ith award is made and $X_i = 0$ otherwise. In such cases, the ith award value is a fixed number ϵ_i (specified by the contract). Then $H_n = \sum_{i=1}^{n} \epsilon_i X_i$.

*It may be that the times of occurrence of milestone events are subject to schedule incentives of the usual incentive fee type as part of the compound contract structure.

†The quantity n is not a random variable. Models for multiple award fee structures in which n is random can be formulated, but are not considered here.

Graded award fee structures employ the random variable X_i to represent the amount of fee awarded at the ith award event. Then $H_n = \sum_{i=1}^{n} X_i$.

Independent awards employ independent random variables X_i (in either the binary or graded case) whereas dependent award structures may in principle employ any specified distribution $F_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$. Independent binary and graded models will be illustrated in the next section. General considerations relating to dependent models will be presented here before introducing a dependent graded model in a subsequent section.†

A dependent graded multiple award fee model that is intended for use as a contract structure must be capable of relatively simple analysis to affect contract realization. It must satisfy basic requirements and be nontrivial in form while achieving such simplicity. These requirements rule out a number of possibilities for the dependency relations between the random variables X_i. The only way to analyze most models that may be stipulated is by monte carlo simulation. This is considered unsatisfactory in the present application for two reasons. Such analyses require computer effort, careful construction and interpretation by specialists, and the possibility of time delays. In many instances of actual contract formulation these requirements are highly undesirable. A more basic objection to such models is that one tends to lose track of the meaning of numbers. The outcome of a particular simulation is difficult (or impossible) to relate to desired motivational factors in a contract structure. For these reasons it was taken as a basic rule that the dependent graded model should contain features that could be directly related to motivational factors and yield to at least partial closed form analyses. This is defined to mean that for the stochastic process used cumulative expected values and variances should be obtainable by means of relatively simple calculations.

For the purposes of developing specific model structures several basic assumptions are employed as guidelines in formulating the dependent graded model:

1) Award fee is not the total fee. It should be an additional increment of fee added to other fees because of specific value produced by the contractor. Such value should be beyond any value otherwise motivated and contracted for within the total contractual agreement. Corollary to this the award amounts should be expressible in terms of the typical monetary unit involved. Thus, if a contract deals with several thousand dollars the award should be in less than thousand amounts. A contract dealing with millions of dollars might employ awards of several thousand dollars, and so forth.

2) Rewards should be propagated. A high award fee at an award event should influence future awards to be greater because of the already demonstrated efforts of the contractor.

3) Penalty should be propagated. A low award fee at an award event should influence awards to be less because of the alrady demonstrated defects of the contractor.

4) The range of possible award fee distributions should not be excessive. The best award situation should still have a sufficient range of possibilities as represented by the structures random character to provide latitude for a contractor to exercise options. Such built-in possibilities for leverage by both parties provides one of the major attractive features of award fee contracts. Similarly the worst award situation, though reducing the probability of higher award fees should still provide for the possibility of such awards.

†Dependent binary models are not discussed further. They are not felt to be of general interest as models for award fee structures.

These four assumptions are taken as the defining properties underlying the Award Fee concept and are therefore used as building blocks for award fee contract structures.

The concept of risk assumption in different contract types is an important factor in selecting types and in formulating the contract structure (as e.g., during contract negotiations). Risk assumption for the nonaward fee contract types has been quantified [11] and extensively discussed (e.g., in [17] and [15]). The risk assumption in Award Fee contracts is of a different kind than risk in the other types. Since one does not deal with a pure Award Fee contract the element of risk due to Award Fee is distinct from whatever risk is otherwise present in a contract situation. The award fee risk is only on extra fee and should not be considered as a major factor of risk assumption by a contractor. There is a special subjective element in such risk due to the a posteriori nature of the award fee. Thus one may argue that award fee risk should not be considered at all as part of the entrepreneural risk assumption, but only as a measure of the variability in total award fee. The cumulative variance of multiple awards then serves very well as a measure of such risk.

INDEPENDENT MODELS

When the multiple awards are represented in terms of independent random variables the contract structure model is simple to formulate and analyze. In the independent case the joint distribution function is the product of the n-marginal distributions associated with the n award events, i.e.:

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)=\prod_{i=1}^{n}F_{X_i}(x_i).$$

In the case of graded independent awards the expected value and variance of H_n, the total award fee are given as follows in terms of X_i = award fee at the ith award event:

$$A_n=E\,[H_n]=\sum_{i=1}^{n}E\,[X_i],$$

$$S_n=\mathrm{Var}\,[H_n]=\sum_{i=1}^{n}\mathrm{Var}\,[X_i].$$

When each of the random variables X_i has the same distribution, extremely simple formulas result:

$$A_n=n\,E\,[X],$$

$$S_n=n\,\mathrm{Var}\,[X],$$

where X denotes the typical award fee random variable. These formulas are useful for comparison with more complex structural models and serve as guides to the possible value of expectation and variance.

For binary awards in which X_i represents the event in which an award of fixed amount ϵ_i is granted or not an independent binary award fee structure results. Let $Pr\,[X_i=1]=p_i$ and $P_r\,[X_i=0]=1-p_i=q_i$. Then:

$$A_n=\sum_{i=1}^{n}\epsilon_i p_i,$$

$$S_n = \sum_{i=1}^{n} \epsilon_i p_i q_i.$$

When $p_i = p$ for $i = 1, \ldots, n$ one obtains the simple results:

$$A_n = pY_n,$$

$$S_n = pqY_n,$$

where Y_n denotes the total award fee pool available (in case each ϵ_i is the same value $Y_n = n\epsilon$, otherwise $Y_n = \sum_{i=1}^{n} \epsilon_i$). The possibilities for structuring independent binary award fee contracts using this model are shown in Figure 1. The possibilities are limited, but their very simplicity makes them important as guidelines and illustrations. The two quantities p and Y_n are directly related to real quantities (likelihood of award and total award fee) and the effects produced by changing them may be readily appreciated in terms of specific contractual goals.

When X_i represents the amount of award fee granted at the ith award event and independent graded awards are employed the above formulas may be applied to analyze such contract structures once the $F_{X_i}(x_i)$ are specified. Formulation of graded structures requires the specification of those distributions and should employ analysis of proposed forms to determine which ones will produce proper contract motivation (in terms of expected return A_n and risk S_n to the contractor). Since the analysis is so simple in the independent case any set of distributions may be studied and employed if found desirable. The use of normal or uniform or triangular distributions suggest themselves because of their relative simplicity and their ability to express specific characteristics that might be utilized in contract

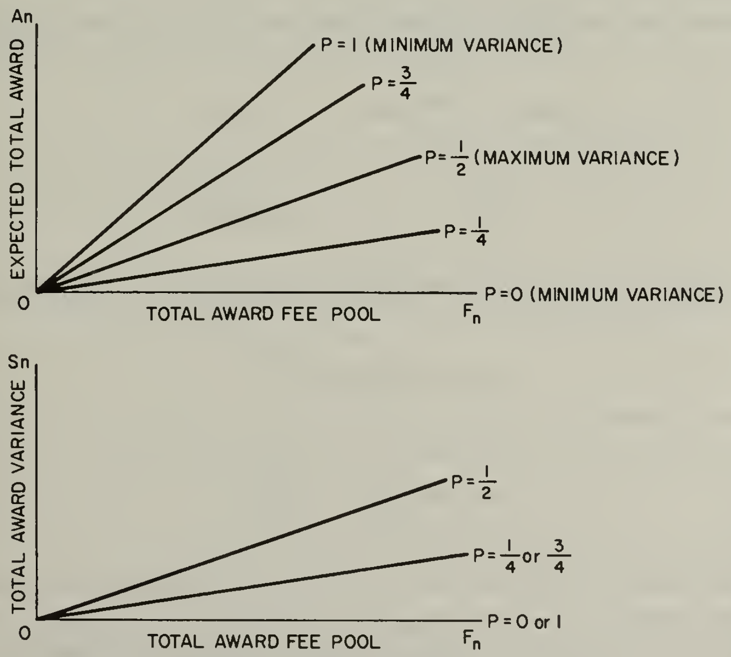

FIGURE 1. Range of results for the binary award fee model.

structure. Thus the uniform distribution would represent a range of award fees without making any more likely than another while a triangular distribution would indicate that some award values are more likely than others.

In the construction of more complex, dependent models one may wish to employ a related independent model for comparison. If such a technique is intended the independent model must be selected in terms of the analysis required for the corresponding nonindependent model. Trapezoidal distributions are relatively simple to work with in constructing and analyzing award fee contract structures and they allow control over range and relative occurance of fee amounts. Such distributions are employed in the dependent model developed in the next section. They will be introduced here and used to formulate a corresponding independent structural model for multiple graded award fee contracts.

A general trapezoidal density function $f(\epsilon)$ is shown in Figure 2. It is specified by two parameters,

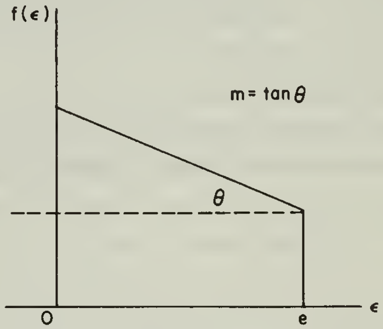

FIGURE 2. Trapezoidal density function.

the range e and the slope $m = \tan\theta$. From Figure 2 one sees that $f(\epsilon)$ has the following expression:

$$f(\epsilon) = \frac{1}{e} + \frac{me}{2} - m\epsilon, \qquad 0 \leq \epsilon \leq e, \tag{1}$$

$$= 0, \qquad \text{elsewhere.}$$

A random variable ϵ specified by the trapezoidal distribution $f(\epsilon)$ has the following statistics:

$$E[\epsilon] = \frac{e}{2}\left(1 - \frac{me^2}{6}\right), \tag{2}$$

$$\text{Var}\,[\epsilon] = \frac{e^2}{12}\left(1 - \frac{m^2e^4}{12}\right). \tag{3}$$

By varying the parameters e and m one may obtain a variety of distributions, the extreme cases are the uniform distribution when $m = 0$ and the triangular distribution (with apex at $\epsilon = 0$) when $m = 2/e^2$. Table 1 shows the ranges of parameter variations that are possible in specifying trapezoidal density factions.

TABLE 1

e	Maximum Value of m (to nearest integer)
0.1	200
0.2	50
0.3	22
0.4	12
0.5	8
0.6	5
0.7	4
0.8	3
0.9	2
1.0	2

A model in which there is indicated a likelihood of obtaining greater award fee results as the form of the distribution approaches the uniform case ($m = 0$). Thus variation in m controls the frequency with which any value of ϵ occurs. In the extreme, uniform case all values become equally likely. Higher values never become more likely in models based on the trapezoidal distribution of Figure 2. Such a situation is desirable in view of the assumed role of ϵ as an increment of total fee awarded in addition to base fees. Major values of fee should be incorporated as part of the nonaward portion of a contract structure in most cases. Of course other distributions may be used for cases in which the trapezoidal form (with maximum at zero) is considered inappropriate. The remainder of this paper will deal only with trapezoidal models in formulating, analyzing, and illustrating multiple award fee structures.

The parameter e specifies the possible range of values that may be assumed by the award fee ϵ. In the independent model total award fee H_n is analyzed in terms of the individual random award fees ϵ_i, $i = 1, \ldots n$ that occur at the ith award events. Each ϵ_i has a trapezoidal density specified by a slope m_i and range e_i. The simplest case is when each ϵ_i is from the same distribution and only two parameters m, and e are required to specify the structure. As discussed for the simple binary model this special case is useful for appreciating the role of parameters and structure forms and for comparison with complex dependent models. The values for $E[\epsilon]$ and $\text{Var}[\epsilon]$ given by Equations (2) and (3) provide all the necessary analysis in the independent case since $A_n = n\,E[\epsilon]$, and $S_n = n\,\text{Var}[\epsilon]$. It may be observed that $E[\epsilon]$ is cubic in e and linear in m whereas $\text{Var}[\epsilon]$ is quintic in e and quadratic in m. Thus both statistics are less sensitive to changes in slope than to changes in range. This is desirable in view of the fact that the range represents a real quantity (fee) directly whereas the slope has a less direct physical interpretation. The slope relates to motivational aspects of the contract in that it controls the likelihood of various fees to some extent. Such relations are somewhat subjective and difficult to specify exactly requiring relative insensitivity of the structure parameter if reasonable models are to be constructed.

A numerical example of the simple independent graded structure will be included (for comparison) with examples of the dependent model in the next section.

DEPENDENT GRADED MODEL

This model is based on the trapezoidal distribution described in the previous section. Dependence is introduced by allowing the slope of each distribution to be a random variable m_i determined by the outcome ϵ_{i-1} of the previous award event. The range values e_i are specified parameters of the award fee contract structure. The dependence specification relation between the random variables m_i and ϵ_{i-1} is assumed to be:

$$m_i = \frac{2}{e_{i-1}e_i^2}\,(e_{i-1} - \epsilon_{i-1}). \tag{4}$$

When $\epsilon_{i-1} = e_{i-1}$ this relation gives $m_i = 0$. Thus when the maximum possible award fee occurs at the $i-1$ award event the next award fee ϵ_i is distributed in the most favorable way possible in the contract structure. This effect may be considered to represent a propagation of reward, good work in the present provides an increased liklihood of higher award fee in the future. Such a situation is often desirable as a reflection of real life efforts and provides desirable motivation within contract structure (of the carrot variety). On the other hand when $\epsilon_{i-1} = 0$, $m_i = 2/e_i^2$, which is the least desirable distribution of award fee from the contractors point of view. The effect is a propagation of penalty which reflects the likelihood of low future awards when low present awards occur. This also provides a desirable form of motivation within contract structure (of the stick variety). Thus the dependence introduced into this model is seen to represent useful practical forms of motivation for award fee contract structures. Moreover, the stochastic process specified by the set of random variables $\{\epsilon_1\}$ lends itself to analysis in relatively simple terms. The values of A_n and S_n may be computed in terms of difference equations that will now be developed by employing techniques of conditional probability distributions.

Let n denote the number of award events, e_i the range of the ith award fee ϵ_i, and M_1 the initial slope. The slopes m_i for $i > 1$ are random variables specified by Equation (4).

The basic formulas are:

Conditional density function

$$f(\epsilon_i \mid m_i) = \frac{1}{e_i} + \frac{m_i e_i}{2} - m_i \epsilon_i, \tag{5}$$

conditional expectation

$$E\,[\epsilon_i \mid m_i] = \frac{e_i}{2}\left(1 - \frac{m_i e_i^2}{6}\right), \tag{6}$$

then,

$$E\,[\epsilon_i] = E\,[E\,[\epsilon_i \mid m_i]] = \frac{e_i}{2} - \frac{e_i^3}{12} E\,[m_i], \tag{7}$$

conditional variance $\operatorname{Var}\,[\epsilon_i \mid m_i] = E\,[(\epsilon_i - E\,[\epsilon_i \mid m_i])^2 \mid m_i]$

$$\operatorname{Var}\,[\epsilon_i \mid m_i] = \frac{e_i^2}{12}\left[1 - \frac{m_i^2 e_i^4}{12}\right] \tag{8}$$

using $\text{Var}[\epsilon_i] = E[\text{Var}[\epsilon_i \mid m_i]] + \text{Var}[E[\epsilon_i \mid m_i]]$ one obtains

$$\text{Var}[\epsilon_i] = \frac{e_i^2}{12}\left[\frac{2}{3} + \frac{2}{3}\frac{E[\epsilon_{i-1}]}{e_{i-1}} - \frac{1}{3}\frac{(E[\epsilon_{i-1}])^2}{e_{i=1}^2}\right]. \tag{9}$$

One may derive the following useful expression related to the expected values (by using the assumed relation (4) between ϵ_i and m_i):

$$E[\epsilon_i] = e_i - E[m_{i+1}]\, e_{i+1}^2\, e_i/2. \tag{10}$$

Let $M_i \equiv E[m_i]$ and $E_i \equiv E[\epsilon_i]$ to simplify notation. Then equating values of E_i in (7) and (10) yields:

$$M_i = \frac{1}{e_i^2} + \frac{e_{i-1}^2}{6e_i^2} M_{i-1}, \qquad i>1 \tag{11}$$

and M_1 is the assumed initial slope of the density of ϵ_1.

Alternatively by solving (7) and (10) for M_i and equating the results one obtains:

$$E_i = \frac{e_i}{3} + \frac{e_i}{6e_{i-1}} E_{i-1}, \qquad i>1 \tag{12}$$

where

$$E_1 = \frac{e_1}{2}\left(1 - \frac{M_1 e_1^2}{6}\right).$$

Solving either of the recurrence relations (11) or (12) yields:

$$E_i = e_i\left(\frac{2}{5} + \left(\frac{1}{6}\right)^i\left(\frac{3}{5} - \frac{e_1^2 M_1}{2}\right)\right). \tag{13}$$

As i, the number of award events becomes large, the expected value of a particular award becomes independent of the initial distribution parameters e_1 and M_1 and tends to the value $2e_i/5$ which falls between the possible extreme values of $e_i/2$ and $e_i/3$ corresponding to the uniform and triangular densities, respectively.

Since Equation (13) gives an explicit expression for E_i, Equation (9) can be used to obtain an explicit expression for $\text{Var}[\epsilon_i]$. It may be observed that as i becomes large $\text{Var}[\epsilon_i]$ tends to an expression independent of the initial distribution of ϵ_1. The explicit expressions for E_i and $\text{Var}[\epsilon_i]$ are not developed further here since one is not particularly interested in these quantities, but rather in expressions for the total award expectation and variance A_n and S_n.

The expected total award fee, A_n is obtained directly from the above results.

$$A_n = E[H_n] = \sum_{i=1}^{n} E[\epsilon_i] = \sum_{i=1}^{n} E_i,$$

using the value of E_i provided by Equation (13) one obtains

$$A_n = \frac{2}{5} \sum_{i=1}^{n} e_i + \left(\frac{3}{5} - \frac{e_1^2 M_1}{2}\right) \sum_{i=1}^{n} e_i \left(\frac{1}{6}\right)^i, \qquad n \geqslant 1. \tag{14}$$

Equation (14) gives an explicit expression for A_n in terms of the contract structure parameters M_1 and e_i, $i=1, \ldots, n$. In many cases it is more useful to employ a recurrence relation to generate the A_n in terms of the previous expected total award A_{n-1} (when one less award event occurs). Such an expression is obtained from Equation (14).

$$A_{n+1} = A_n + \frac{2}{5} e_{n+1} + \left(\frac{3}{5} - \frac{e_1^2 M_1}{2}\right) e_{n+1} \left(\frac{1}{6}\right)^{n+1}, \qquad n > 1. \tag{15}$$

When the parameters e_i all equal the same value, say e, then a simple explicit formula is obtained for A_n:

$$A_n = \frac{e}{5}\left[2n + \left(\frac{3}{5} - \frac{e^2 M_1}{2}\right)\left(6 - \left(\frac{1}{6}\right)^n\right)\right]. \tag{16}$$

As an alternative to the use of the recurrence relation (15) one may employ two simple recurrence relations for A_n and E_n:

$$A_{n+1} = A_n + E_{n+1} \tag{17}$$

and

$$E_{n+1} = \frac{e_{n+1}}{3} + \frac{e_{n+1}}{6e_n} E_n. \tag{18}$$

One uses (18) to generate E_{n+1} then substitutes this into (17) to generate the next A_{n+1} value.

Previously it has been remarked that it is desirable for the analysis of award fee structures to have expressions for the variance of total award fee as well as for its expectation. Such expressions represent measures of contractor risk and may be used as quantitative features of a contract. Expressions for $S_n = \mathrm{Var}\,[H_n]$ will now be obtained though they are of course more complex than the relatively simple results obtained above for A_n.

$$S_{n+1} = \mathrm{Var}\,[H_{n+1}] = \mathrm{Var}\,[H_n + \epsilon_{n+1}],$$

$$S_{n+1} = \mathrm{Var}\,[H_n] + \mathrm{Var}\,[\epsilon_{n+1}] + 2\{E[H_n \epsilon_{n+1}] - E[H_n]E[\epsilon_{n+1}]\},$$

$$S_{n+1} = S_n + \mathrm{Var}\,[\epsilon_{n+1}] + 2[R_{n+1} - A_n E_{n+1}]. \tag{19}$$

Equation (19) would yield a recurrence for S_n in terms of the previously obtained quantities $\text{Var}[\epsilon_{n+1}]$, A_n, and E_{n+1} provided the quantity $R_{n+1} \equiv E[H_n\epsilon_{n+1}]$ was known. Now,

$$R_{n+1} = E\left[\sum_{i=1}^{n} \epsilon_i\epsilon_{n+1}\right] = \sum_{i=1}^{n} E[\epsilon_i\epsilon_{n+1}],$$

thus we are lead to study the quantities $E[\epsilon_i\epsilon_{n+1}]$, $i = 1, \ldots, n$.

Consider $E[\epsilon_i\epsilon_k] = E[E[\epsilon_i\epsilon_k|m_k]]$, for $i = 1, \ldots, k-1$.

$E[\epsilon_i\epsilon_k|m_k] = \epsilon_i E[\epsilon_k|m_k]$ since ϵ_i is independent of m_k. Thus the conditional expectation is

$$E[\epsilon_i\epsilon_k|m_k] = \epsilon_i \frac{e_k}{2}\left(1 - \frac{m_k e_k^2}{6}\right).$$

Using (4) to specify the random variable m_k in terms of ϵ_{k-1}, one obtains

$$\frac{m_k e_k^2}{6} = \frac{1}{3} - \frac{\epsilon_{k-1}}{3e_{k-1}},$$

so that

$$E[\epsilon_i\epsilon_k|m_k] = \epsilon_i e_k \left(\frac{1}{3} + \frac{\epsilon_{k-1}}{6e_{k-1}}\right),$$

and

$$E[\epsilon_i\epsilon_k] = \frac{e_k}{3} E_i + \frac{e_k}{6e_{k-1}} E[\epsilon_i\epsilon_{k-1}]. \tag{20}$$

Equation (20) gives a form of recurrence relationship for the quantities $E[\epsilon_i\epsilon_k]$. Using this relation for $i = 1, \ldots, k-1$ one obtains:

$$R_k \equiv \sum_{i=1}^{k-1} E[\epsilon_i\epsilon_k],$$

$$R_k = \frac{e_k}{3} \sum_{i=1}^{k-1} E_i + \frac{e_k}{6e_{k-1}} \sum_{i=1}^{k-2} E[\epsilon_i\epsilon_{k-1}] + \frac{e_k}{6e_{k-1}} E[\epsilon_{k-1}^2],$$

thus

$$R_k = \frac{e_k}{3} A_{k-1} + \frac{e_k}{6e_{k-1}} R_{k-1} + \frac{e_k}{6e_{k-1}} E[\epsilon_{k-1}^2], \tag{21}$$

where

$$E[\epsilon_{k-1}^2] = \frac{1}{6} e_{k-1}^2 + \frac{e_{k-1}^2}{6e_{k-2}} E_{k-2}.$$

By using the results obtained above one has the following recurrence relation for R_n

$$R_{n+1} = \frac{e_{n+1}A_n}{3} + \frac{e_{n+1}R_n}{6e_n} + \frac{e_{n+1}e_n}{36} + \frac{e_{n+1}e_n}{36e_{n-1}} E_{n-1}. \tag{22}$$

One uses this expression to obtain R_{n+1} which is then substituted in (19) to obtain the next value for S_{n+1}. In this way the variance expressions S_n are obtained recursively. It may be remarked that the recurrence technique is particularly appropriate for analysis of the multiple award fee contracts since most structures deal with only a few (under 10) award events.

The use of trapezoidal distributions as the basis for a model of multiple graded award fee contracts leads to the analysis presented above. By specifying the number of award events n, the slope M_1 of the initial density, and the range parameters e_i for $i=1, \ldots, n$ a great variety of award fee contract structures can be obtained. One can utilize the evaluation of A_n and S_n to analyze different structures and select those that most closely represent the buyer's wishes. A contractor too may carry out its analyses as part of the contract negotiation procedure or to arrive at a prenegotiations-position.

The wide variety of possibilities makes presentation of extensive numerical results unfeasible. However, some insight can be gained into the possibilities for dependent structuring by considering some extreme results for the case in which all the range parameters are equal. Such structures are specified by two parameters, the initial slope M_1, and the common range value e, In each of the examples five award events were considered and the results were compared with the corresponding independent case for total (five award events) award. Tables 2 and 3 show the results for range parameter $e=0.1$ for initial slope values of zero and of 100 respectively. Tables 4 and 5 show the results for range parameter $e=1$ for initial slope values of zero and of 2 respectively. As specified in Table 1 these values for initial slopes represent values within the region of possible slope values. The independent model results utilize the same trapezoidal density for each award event, each density is specified by the same two parameters e and M_1.

TABLE 2

($e=0.1$, $M_1=0.0$)

n	A_n	Independent model value	S_n	Independent model value
1	0.05		0.001	
2	0.09		0.002	
3	0.13		0.003	
4	0.17		0.004	
5	0.21	0.25	0.005	0.005

TABLE 3

($e=0.1$, $M_1=100$*)

n	A_n	Independent model value	S_n	Independent model value
1	0.04		0.001	
2	0.08		0.002	
3	0.12		0.003	
4	0.16		0.004	
5	0.20	0.16	0.005	0.005

*$M=200$ is the maximum possible slope (to nearest integer) when $e=0.1$, the value $M_1=100$ indicates the kind of results obtained, for comparison with $M_1=0$.

TABLE 4

($e=1.0,\quad M_1=0.0$)

n	A_n	Independent model value	S_n	Independent model value
1	0.50		0.083	
2	0.92		0.188	
3	1.32		0.291	
4	1.72		0.395	
5	2.12	2.50	0.497	0.415

TABLE 5

($e=1.0,\quad M_1=2^*$)

n	A_n	Independent model value	S_n	Independent model value
1	0.33		0.056	
2	0.72		0.146	
3	1.12		0.245	
4	1.52		0.347	
5	1.92	1.65	0.449	0.280

*$M_1=2$ is the maximum possible slope (to the nearest integer) when $e=1.0$.

Several observations can be made from the results presented in Tables 2 through 5 that serve to check on the formalism of the model and provide some insight into the construction of appropriate contract structure.

In each of the four cases the relative insensitivity to initial slope values is clearly demonstrated. However, there is a very definite initial slope effect particularly on the high range ($e=1.0$) expected total values. Variance is less affected by M_1 values. The effect of changing range values is very strong, affecting both expected award and variance, as seen by comparing the results for $e=0.1$ with results for $e=1.0$.

The cases in which $M_1=0$ corresponds to starting with the best possible award fee distribution from the point of view of the contractor. In these cases the independent model gives a higher expected total award fee as one would expect. The possibility of increased (random) slope values on events following the initial award event produces a decrease in the expected total award in the dependent model. The opposite effect is seen for the high initial slope cases. In these the tendency toward low award values imposed by the high initial slope is maintained over all award events in the independent model. The dependent model tends to reduce this effect by introducing a likelihood of improved award values.

For the small range value $e=0.1$ there is little variance in any of the award events. Total variance is essentially the same in both the independent and dependent model, being on the order of 5 percent of the range value e. However, in the higher range $e=1.0$ cases there are significant values of variance. Total variance reaches on the order of fifty percent of e in the dependent model. Considerable difference exists in variance between the independent and dependent models, this is particularly true at

the high slope value. In each case the independent model gives less variance than the value of the dependent model as one would expect.

As a final reference to the numerical results one might remark on the fact that the independent and dependent models do not differ by what seems to be significant values. However one must place those values in the context of a contract structure application.

In such a context for example the total expected dependent award of 1.92 and the total expected independent award of 1.65 in the $e = 1.0$, $M_1 = 2.0$ case (Table 5) represents a difference of $270,000.00 when the unit of contract value is a million dollars.

The results presented and described above indicate the wide variety of possible award fee contract structures that can be developed by considering tradeoffs between expected total returns and variability. In general, some combinations of e and M_1 can be obtained that will yield a satisfactory combination from the viewpoint of both parties. The consideration of such structures can be assisted by employing independent model examples and indeed the independent model may be the most desirable in some cases. It should be noted, however, that the dependent model incorporates a greater possibility of control by the buyer over contractor performance after the contract is signed. These follow from the propagation of rewards and of penalties that are inherent in the dependent model. It is felt that this model provides a realistic method for incorporating controls into the mathematical contract structure.

REFERENCES

[1] Condon, J. E., "Mathematical Models for Multiple Incentive Contracting," doctoral dissertation, George Washington University, 1967.

[2] Cook, J. and others, "The Evaluation and Structuring Techniques of Multiple Incentive Contracts," Dept. of the Air Force, 1966.

[3] Cravens, J. E., "The Challenges of Cost Plus Award Fee Contracting." Remarks at Defense Procurement Executive Refresher Course, March 1968.

[4] Cravens, J. E., and Rule, G. W., "The Past and Future in Cost Plus Award Fee Contracting," discussing at Materiel Secretaries Breakfast Meeting, October 1968.

[5] Fisher, I. N., "A Reappraisal of Incentive Contracting Experience," Rand Memo RM–5700–PR, July 1968; and "Improving and Effectiveness of Incentive Contracting," Rand Paper P–3870, July 1968.

[6] Hagen, W. A., "Systematic Technique of Incentive Contracting (STOIC)," Marshall Space Flight Center NASA, 1966.

[7] Jones, T. H., "A Study of Selected Aspects of the Use of Incentive Contracts in U.S. Air Force Procurement Management," doctoral dissertation, Ohio State University, 1967.

[8] LMI, "An Examination of the Foundations of Incentive Contracting," Logistics Management Institute, May 1968.

[9] Marshall, C. W., "Proposal Evaluation Technique Conditioned On Contract Kind (PETCOCK) Report," Polytechnic Institute of Brooklyn, 1967.

[10] Marshall, C. W. and Ahmadi, H., "Contributions to Multiple Incentive Contracting Report," Polytechnic Institute of Brooklyn, 1967.

[11] Marshall, C. W., "Quantification of Contractor Risk," Nav. Res. Log. Quart., *16*, 531–541 (1969).

[12] NASA/DOD *1969 Incentive Contracting Guide.*

[13] NASA, *Cost Plus Award Fee Contracting Guide*, 1967.

[14] Peck, M. J. and Scherer, F. M., *The Weapons Acquisition Process: An Economic Analysis* (Harvard U. Press, 1962).

[15] Pratt, J. W., "Risk Aversion in the Small and in the Large," Econometrica, *32*, 132–136 (Jan.–Apr. 1964).

[16] Scherer, F. M., *The Weapons Acquisition Process: Economic Incentives* (Harvard U. Press, 1964).

[17] Seagle, J. P., "Risk Aversion in Contract Negotiations," paper presented at TIMS/ORSA meeting, May 1968.

ORDER-PRESERVING ALLOCATION OF JOBS TO TWO MACHINES

S. Mehta

R. Chandrasekaran

H. Emmons

Case Western Reserve University

Cleveland, Ohio

ABSTRACT

In this paper, we consider the problem of minimizing the mean flow time of jobs to be processed on two machines. The jobs have a predetermined order, perhaps reflecting the order of arrival, and each job has a known processing time. We wish to assign the jobs to machines so as to minimize the mean flow time, with the constraint that the original order must be preserved within the subset of jobs assigned to each machine. An efficient algorithm based on dynamic programming is developed.

INTRODUCTION

Sequencing problems in which the objective is to minimize the mean flow time of the jobs are quite common in the literature. (See [1].) Problems of scheduling n jobs on m identical machines have been considered [1]. In all of these, the jobs can be reordered in any manner the scheduler decides; and this works as long as the jobs are either inanimate, such as jobs in a machine shop, or if one does not explicitly see the queue as in a computer. However, when the objects being sequenced are human beings, as in the case of patients in a hospital emergency clinic or customers in a supermarket, the situation changes. In these cases we have to preserve, in some manner, the order in which they arrive at the service center. It is this kind of a problem that we consider here. In this paper, we restrict our attention to the case of two machines.

Suppose each of n jobs are to be processed on either of two identical machines. The jobs have an arbitrary predetermined order, perhaps reflecting their order of arrival, and each job has a known processing time. We wish to assign jobs to machines so as to minimize the mean flow time (or equivalently total flow time), with the constraint that the original order must be preserved within each of the two subsets of jobs assigned to each machine. In case the processing times are not known precisely in advance, it is easy to see that their expected values may be used to obtain the minimum total (or mean) expected flow time. For example, a hospital emergency clinic with two physicians may wish to form n waiting patients into two queues so as to speed the flow. Each patient may be arbitrarily assigned to either physician, but is unwilling to allow a later arrival to precede him.

THE ALGORITHM

Let us index the jobs in reverse order, so that the last arrival is job 1, etc., and assume that job j ($j=1, \ldots, n$) has processing time t_j. The following dynamic programming approach solves the problem efficiently.

Let $f_j(k)$ be the minimum value of total flow times of jobs $1, \ldots, j$ obtained by putting job j in position k from the end (see Figure 1) in one of the queues $(1 \leqslant k \leqslant j)$. Note that the increase in the value of objective function incurred by putting job j ahead of $k-1$ other jobs does not depend on which jobs are behind job j; only on how many. If job j goes to position k from the end, clearly $j-1$ must have gone to either position $k-1$ in the same quene or position $j-k$ in the other quene. This gives us the recurrence relation:

$$f_j(k) = kt_j + \min\{f_{j-1}(k-1), f_{j-1}(j-k)\}, \tag{1}$$

with $f_1(1) = t_1$ and $f_j(0) = \infty$ $(j = 1, \ldots, n)$.

Thus, at each stage $j\,(j = 1, \ldots, n)$ (1) gives $f_j(k)\,(k = 1, \ldots, j)$ in terms of the previous stage results.

The following identity can be used to simplify calculations.

$$f_j(j-k+1) = (j-2k+1)t_j + f_j(k). \tag{2}$$

This follows immediately, using (1) to evaluate $f_j(j-k+1) - f_j(k)$. Using (2), with j set to $j-1$, we can rewrite (1):

$$f_j(k) = kt_j + \min\{f_{j-1}(k-1), (j-2k)t_{j-1} + f_{j-1}(k)\} \tag{3}$$

with $f_1(1) = t_1$ and $f_j(0) = f_j(j+1) = \infty$ $(j = 1, \ldots, n)$.

In addition, note that for $k = (j+1)/2$ (j odd), (1) gives

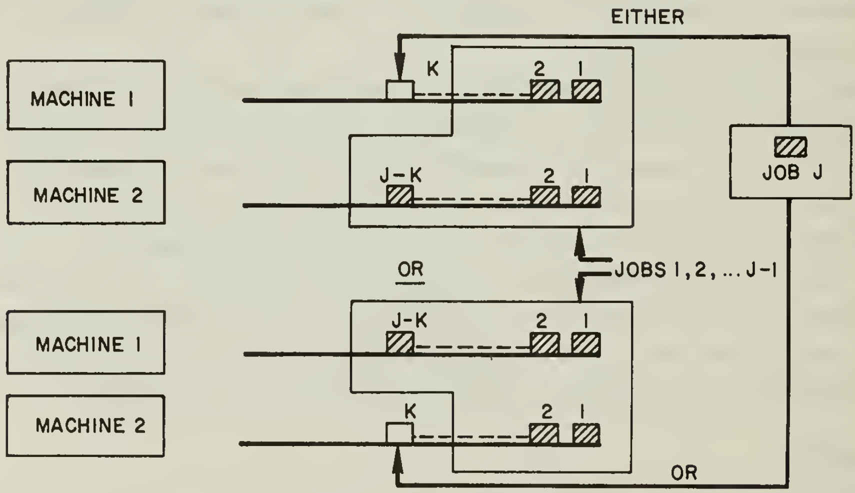

FIGURE 1

$$f_j(k) = kt_j + f_{j-1}(k-1) \tag{4}$$

We now observe that at each stage we need only compute $f_j(k)$ for half the values of k; specifically, for:

$$1 \leqslant k \leqslant \left\{ \begin{array}{l} j/2,\ j \text{ even} \\ (j+1)/2,\ j \text{ odd} \end{array} \right\} \equiv \left[\frac{j+1}{2} \right], \qquad \text{or } 1 \leqslant k < \frac{j}{2}+1.$$

Note, (2) implies that for $k \geqslant j/2+1$,

$$j-2k+1 \leqslant -1$$

$$(j-2k+1)t_j < 0$$

$$f_j(k) > f_j(j-k+1).$$

Thus, at the last stage $(j=n)$, we need only investigate $k < n/2+1$ in order to find the minimizing schedule. Second, using (3) and (4) it can easily be seen that, to compute $f_j(k)$ for $k=1, \ldots, [(j+1)/2]$, knowledge of $f_{j-1}(k)$ for $k=1, \ldots, [(j-1)+1/2]$ is enough so that for $j=2, \ldots, n$ this is a sufficient range for k.

The total number of calculations required in using this algorithm can be estimated by noting that, at stage j there are roughly $j/2$ comparisons and $3j/2$ additions, for a total of $n(n+1)/4$ comparisons and $3n(n+1)/4$ additions. Thus, the algorithm has polynomial (square) bound. It gives all optimal solutions, if more than one exists.

EXAMPLE:

Suppose there are $n=5$ jobs, with

jobs	5	4	3	2	1
processing times	1000	3	1	800	1001

STAGE 1:

$$f_1(1) = t_1 = 1001.$$

STAGE 2:

$$f_2(1) = t_2 + \min\{f_1(0), f_1(1)\} = t_2 + f_1(1)$$

$$= 800 + 1001 = 1801.$$

STAGE 3:

$$f_3(1) = t_3 + t_2 + f_2(1)$$

$$= 1 + 800 + 1801 = 2602$$

$$f_3(2) = 2t_3 + f_2(1) \qquad \text{(using (4))}$$

$$= 2 + 1801 = 1803.$$

STAGE 4:

$$f_4(1) = t_4 + 2t_3 + f_3(1)$$

$$= 3 + 2 + 2602 = 2607$$

$$f_4(2) = 2t_4 + \min\{f_3(1), f_3(2)\}$$

$$= 6 + 1803 = 1809.$$

STAGE 5:

$$f_5(1) = t_5 + 3t_4 + f_4(1)$$

$$= 1000 + 9 + 2607 = 3616$$

$$f_5(2) = 2t_5 + \min\{f_4(1), t_4 + f_4(2)\}$$

$$= 2000 + \min\{2607, 3 + 1809\} = 3812$$

$$f_5(3) = 3t_5 + f_4(2)$$

$$= 3000 + 1809 = 4809.$$

Thus, the optimal value of the total flow time is $\min_k f_n(k) = 3616$, with mean flow time $3616/5 = 723.2$. Tracing back, we find the optimal assignment is to put job 5 on one machine, and all other jobs on the other machine.

REFERENCES

[1] Conway, R., W. Maxwell, and L. Miller, *Theory of Scheduling* (Addison-Wesley, 1967).

A NOTE ON THE COSTLY SURVEILLANCE OF A STOCHASTIC SYSTEM*

S. S. Chitgopekar

University of Wisconsin
Madison, Wisconsin

ABSTRACT

We consider the costly surveillance of a stochastic system with a finite state space and a finite number of actions in each state. There is a positive cost of observing the system and the system earns at a rate depending on the state of the system and the action taken. A policy for controlling such a system specifies the action to be taken and the time to the next observation, both possibly random and depending on the past history of the system. A form of the long range average income is the criterion for comparing different policies. If $\mathbf{R}^{\Delta}$ denotes the class of policies for which the times between successive observations of the system are random variables with cumulative distribution functions on $[0, \Delta]$, $\Delta < \infty$, we show that there exists a nonrandomized stationary policy that is optimal in $\mathbf{R}^{\Delta}$. Furthermore, for sufficiently large Δ, this optimal policy is independent of Δ.

1. INTRODUCTION

We are interested in a stochastic system which at any time can be in one of a finite number of L states. There is a positive cost c_0 to observe the system and determine the state it is in. In each of the L states, we have $K(<\infty)$ alternative actions available to us. In state i, it costs $c_i^k \geqslant 0$ to take action k and then the system earns at the rate a_i^k per unit of time, $k=1, \ldots, K$ $i=1, \ldots, L$. The probability of transition from state i to state j, under action k, *if there are no further actions before a transition*, is denoted by p_{ij}^k. We have

$$
(1.1)\qquad \begin{aligned} p_{ij}^k &\geqslant 0, &&\text{for all } i, j, k \\ \sum_{j=1}^{L} p_{ij}^k &= 1, &&\text{for all } i, k. \end{aligned}
$$

When action k is taken in state i, T_i^k denotes the (random) waiting time to transition, *if there are no further actions before a transition*. Let $F_i^k(\cdot)$ be the cumulative distribution function of T_i^k. We assume that $F_i^k(\cdot)$ is continuous, and further, that

$$F_i^k(0)=0; \qquad F_i^k(x)>0 \qquad \text{for all } x>0;$$

$$\left.\frac{d}{dx} F_i^k(x)\right|_{x=0} \text{ exists and is finite;}$$

* This research was supported by the Office of Naval Research under Contract NONR 988(13): NR 042-003, at the Florida State University, Tallahassee, Florida

and

$$E(T_i^k) = \int_0^\infty x dF_i^k(x) = \theta(i, k) < \infty; \qquad k = 1, \ldots, K; \qquad i = 1, \ldots, L.$$

We also assume that in each state i, $i = 1, \ldots, L$, there is at least one action k_i such that the resulting transition probability matrix $P = (p_{ij}^{k_i})$ is the transition probability matrix of an irreducible Markov chain. (See appendix of Chitgopekar [1].)

Chitgopekar [1] studies a similar system when there is no cost of observing the system. As such, the system is continuously observed, the transitions noted, and actions taken then. For the system being studied in this paper, there is a positive cost of an observation, and hence, the transitions of the system cannot be noted.

Any policy S for controlling the system has to specify the times, possibly random, for observing the system, and the action to be taken following an observation, possibly randomized and depending on the past history of the system. If $G(\cdot)$ is a distribution function on $[0, \infty]$ such that $G(t)$ denotes the probability that the system would be observed before t units of time have elapsed since the previous observation, $G(\cdot)$ will be called an *observation distribution.* We restrict our attention to the class of policies $\mathbf{R}^\Delta$ such that for any policy S in $\mathbf{R}^\Delta$ the observation distributions specified by S are distributions on $[0, \Delta]$, $\Delta < \infty$. Any policy S in $\mathbf{R}^\Delta$ specifies, after each observation, the action to be taken and the time to the next observation, both possibly randomized and depending on the past history of the system. The waiting times to transitions, T_i^k, are assumed to be independent of the observation times.

Another assumption which we now make is that if, after an action has been taken following an observation, the system makes a transition before the next observation, the system is temporarily absorbed in the new state and earns nothing until the next observation and action.

As a simple example of the model under study, consider a complex assembly which can be in one of two states: State 1, working; State 2, failed. From State 1(2), the system moves to State 2(1) with probability 1. When the assembly fails, (i.e., a transistion from State 1 to State 2 takes place) unless an observation detects the failure, the system remains in the failed state and earns nothing. The actions available in State 2 are different types of repair service. When the repairs (which take a random time) are complete, the system moves from State 2 to State 1, but will not start production unless an observation determines that repairs are complete and an action is taken to start production.

For any policy S in $\mathbf{R}^\Delta$, the criterion of interest is

$$I(S) = \liminf_{N \to \infty} I_N(S) = \liminf_{N \to \infty} \frac{\sum_{n=1}^{N} Ei_n(S)}{\sum_{n=1}^{N} EY_n(S)}, \tag{1.2}$$

where $Y_n(S)$ is the time from the nth to the $(n+1)$th observation and $i_n(S)$ is the income earned during this period. Let

$$I^* = \sup_{S \in \mathbf{R}^\Delta} I(S). \tag{1.3}$$

We are interested in the existence and nature of policies S^* such that

$$I^* = I(S^*). \tag{1.4}$$

A policy S^* satisfying (1.4) is said to be *optimal.*

DEFINITION: An observation distribution is said to be a *one-point distribution* if there exists an x, $0 \leq x \leq \infty$, such that

$$G(t) = \begin{cases} 0 & \text{for } t < x \\ 1 & \text{for } t \geq x. \end{cases}$$

We shall denote such a distribution by $G_x(\cdot)$.

DEFINITION: A policy S is said to be *nonrandomized stationary* if for each i, $i = 1, \ldots, L$, S specifies a pair $(k_i, G_{x_i}(\cdot))$ such that, whenever the system is observed in state i, S prescribes the action k_i and the one-point observation distribution $G_{x_i}(\cdot)$.

Let $\mathbf{R}_0 \subset \mathbf{R}^\Delta$ be the subclass of nonrandomized stationary policies. Our main result is

THEOREM 1.1: *There exists a nonrandomized stationary policy* S^* *that is optimal in* $\mathbf{R}^\Delta$. *Furthermore, for sufficiently large* Δ, S^* *is independent of* Δ.

The proof of this theorem will be given in section 3.

2. SOME PRELIMINARY FORMULAS AND RESULTS

Suppose under a policy S, at some stage we are in state i, action k is taken, and the next observation is planned after time Y whose distribution function is $G(\cdot)$. Then let

$$\eta(i, k; G) = E \min (T_i^k, Y) = \int_0^\Delta \left\{ \int_0^Y (1 - F_i^k(t))\, dt \right\} dG(y). \tag{2.1}$$

$$\eta(i, k; \Delta) = E \min (T_i^k, \Delta) = \int_0^\Delta (1 - F_i^k(t))\, dt. \tag{2.2}$$

$$q(i, k; G) = Pr(T_i^k \leq Y) = \int_0^\Delta F_i^k(y)\, dG(y). \tag{2.3}$$

(2.4) $p_{ij}'^k$ = the probability that the system will be in state j at the time of the next observation

$$= p_{ij}^k Pr(T_i^k \leq Y) + \delta_{ij} Pr(T_i^k > Y)$$

$$= p_{ij}^k q(i, k; G) + \delta_{ij}(1 - q(i, k; G)),$$

where

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

The expected income before the next observation will be $a_i^k \eta(i, k; G) - c_i^k - c_0$.

Let

$$\theta^* = \max_{i,k} \theta(i, k); \; \theta_* = \min_{i,k} \theta(i, k);$$

$$\eta^*(\Delta) = \max_{i,k} \eta(i, k; \Delta); \; \eta_*(\Delta) = \min_{i,k} (i, k; \Delta);$$

$$a^* = \max_{i,k} a_i^k; \; a_* = \min_{i,k} a_i^k;$$

$$c^* = \max_{i,k} c_1^k; \; c_* = \min_{i,k} c_1^k;$$

and

$$\mu_0 = \begin{cases} \Delta(c_0 + c_* - a^* \eta_*(\Delta)) / (c^* - a_* \eta^*(\Delta)) & \text{if } a^* < 0 \\ (c_0 + c_*) \Big/ \left(a^* + \dfrac{c_0 + c^*}{\Delta} \right) & \text{if } a_* > 0 \\ (c_0 + c_*) \Big/ \left(a^* + \dfrac{c^* + c_0 - a_* \eta^*(\Delta)}{\Delta} \right) & \text{if } a_* < 0 < a^*. \end{cases}$$

We then have

LEMMA 2.1: *Any policy* S *for which* $\liminf_{N \to \infty} \left\{ \sum_{n=1}^{N} \mathrm{E}Y_n(S)/N \right\} < \mu_0$ *is not optimal.*

PROOF: Let S' be a policy that observes the system after every Δ units of time. Then

$$I_N(S') = \sum_{n=1}^{N} E i_n(S') / N\Delta$$

$$\geqslant \begin{cases} -(c_0 + c^*)/\Delta & \text{if } a_* \geqslant 0, \\ (a_* \eta^*(\Delta) - c^* - c_0)/\Delta & \text{if } a_* < 0. \end{cases}$$

Thus

$$I(S') \geqslant \begin{cases} -(c_0 + c^*)/\Delta & \text{if } a_* \geqslant 0, \\ (a_* \eta^*(\Delta) - c^* - c_0)/\Delta & \text{if } a_* < 0. \end{cases} \tag{2.5}$$

Now, for any policy S,

$$I_N(S) = \sum_{n=1}^{N} E i_n(S) \Big/ \sum_{n=1}^{N} E Y_n(S)$$

$$\leqslant \begin{cases} a^*-(c_0+c^*)\Big/\left(\sum_{n=1}^{N} EY_n(S)\Big/N\right) & \text{if } a^* \geqslant 0, \\ (a^*\eta_*(\Delta)-c_0-c_*)\Big/\left(\sum_{n=1}^{N} EY_n(S)\Big/N\right) & \text{if } a^* < 0. \end{cases}$$

Hence

$$I(S) \leqslant \begin{cases} a^*-(c_0+c_*)/\liminf_{N\to\infty}\left\{\sum_{n=1}^{N} EY_n(S)\Big/N\right\} & \text{if } a^* \geqslant 0, \\ a^*\eta_*(\Delta)-c_0-c_*)/\liminf_{N\to\infty}\left\{\sum_{n=1}^{N} EY_n(S)\Big/N\right\} & \text{if } a^* < 0. \end{cases} \tag{2.6}$$

Now, if $\liminf_{N\to\infty}\left\{\sum_{n=1}^{N} EY_n(S)\Big/N\right\} < \mu_0$, from (2.5) and (2.6), we get

$$I(S) < I(S').$$

Hence the Lemma.

Let $\epsilon > 0$ be given. Let $N_{\epsilon,\Delta}$ be a positive integer such that $(N_{\epsilon,\Delta}-1)\epsilon < \Delta \leqslant N_{\epsilon,\Delta}\epsilon$. Take

$$t_i = i\epsilon \qquad i=0, 1, \ldots, N_{\epsilon,\Delta}-1,$$

$$t_{N_{\epsilon,\Delta}} = \Delta.$$

Then from Theorem 2.2 of Chitgopekar [1], it follows that there exists a discrete distribution $G_*(\cdot)$ with its mass only at the points t_i, $i=0, 1, \ldots, N_{\epsilon,\Delta}$, such that

$$q(i,k;G) = \int_0^\Delta F_i^k(t)dG(t) = \int_0^\Delta F_i^k(t)dG_*(t) = q(i,k;G_*),$$

$$|\eta(i,k;G)-\eta(i,k;G_*)| \leqslant 2\epsilon \qquad \text{and}$$

$$\left|\int_0^\Delta t\,dG(t) - \int_0^\Delta t\,dG_*(t)\right| \leqslant 2\epsilon.$$

Let $\mathbf{G}_\epsilon$ denote the finite class of distributions $\{G_{t_i}(\cdot) : i=0, 1, \ldots, N_{\epsilon,\Delta}\}$.

DEFINITION: A policy S_ϵ will be said to be an ϵ-approximation of a policy S if $|I(S)-I(S_\epsilon)| \leqslant \epsilon$.

LEMMA 2.2: *For any policy* S *and for any sufficiently small* $\epsilon > 0$, *there exists a policy* S_ϵ, *an* ϵ*-approximation of* S, *such that* S_ϵ *prescribes observation distributions that are randomizations over the class* G_{ϵ^2}.

The proof is similar to the proof of Theorem 4.1 of Chitgopekar [1].

3. MAIN RESULTS

Let $\mathbf{R}_0^*$ be $\mathbf{R}_0$ augmented by policies with arbitrary observation distributions, that is, not restricted to one-point distributions. Theorem 3.1 below will show that this added generality is not advantageous. Let $X_n(S)$ denote the state of the system at the time of the nth observation, $n=1, 2, \ldots$. It is easily seen that $\{X_n(S)\colon n=1, 2, \ldots,\}$ is a Markov chain with transition probability matrix $P'=(p'^k_{ij})$ where p'^k_{ij} is given by 2.4. As in Chitgopekar [1], we can restrict our attention to nonrandomized stationary policies S for which the Markov chains $\{X_n(S)\colon n=1, 2, \ldots\}$ have only one positive class. We then get (Chung [2])

$$I(S) = \frac{\sum_{i=1}^{L} \{\Pi_i \, (a_i^k \eta(i, k; G_i) - c_i^k - c_0)/q(i, k; G_i)\}}{\sum_{i=1}^{L} \{\Pi_i \, E(Y_i^k)/q(i, k; G_i)\}}, \tag{3.1}$$

where $\{\Pi_i\}$ are the stationary probabilities associated with $P=(p_{ij}^k)$. Note that for any nonrandomized stationary policy S, the state i determines the action k in state i and hence, when dealing with a nonrandomized stationary policy S, we can drop the superscripts k.

THEOREM 3.1: Any optimal policy in $\mathbf{R}_0^*$ is in fact in $\mathbf{R}_0$.

The proof is similar to the proof of Theorem 3.2 of Chitgopekar [1]. However, the x_i introduced in Equation 3.8 of Chitgopekar [1] are now defined by

$$V_i(x_i) = \max_{0 \le x \le \Delta} V_i(x), \; i=1, \ldots, L, \tag{3.2}$$

where

$$V_i(x) = (a_i \eta(i; x) - c_i - c_0 - Ix)/F_i(x), \; i=1, \ldots, L. \tag{3.3}$$

THEOREM 3.2: *There exists an optimal policy in the class* $\mathbf{R}_0$.

The proof is essentially the same as the proof of Theorem 3.3 of Chitgopekar [1].

We are now ready to give the:

PROOF OF THEOREM 1.1: The proof of the existence of a policy $S^* \epsilon \mathbf{R}_0$ which is optimal in $\mathbf{R}_\Delta$ is similar to the proof of Theorem 1.1 of Chitgopekar [1]. To indicate the possible dependence of this optimal policy on Δ, we will denote it by S_Δ^* and let $I(S_\Delta^*)=I_\Delta^*$.

Consider a sequence of nonrandomized stationary policies $\{S_n\}$ such that S_n is in $\mathbf{R}^{\Delta_n}$ and in some state i, it involves the use of the observation distribution $G_{\Delta_n}(\cdot)$, $n=1, 2, \ldots$. From (3.1) we see that $I(S_n) \to 0$ as $\Delta_n \to \infty$. In view of this, we shall assume that, for sufficiently large Δ, $I_\Delta^*=I(S_\Delta^*)>0$.

Next, observe that if $G_{x_i}(\cdot)$ are the observation distributions involved in S_Δ^*, then the x_i maximize

$$V_i(x) = (a_i \eta(i; x) - c_i - c_0 - I_\Delta^* x)/F_i(x), \qquad i=1, \ldots, L.$$

To indicate the possible dependence of x_i on Δ, we shall denote them by $x_i(\Delta)$. For some sufficiently large Δ_0, let

$$A_i = \{x_i(\Delta) : \Delta \geqslant \Delta_0\},$$

$$x_i^* = \sup\{x : x\epsilon A_i\}, \qquad i = 1, \ldots, L,$$

and

$$x^* = \max_i x_i^*.$$

Since, for any $\Delta < \infty$, $I_\Delta^* \leq a^*$ and $V_i(x) \to -\infty$ as $x \to \infty$, it follows that $x_i^* < \infty$, $i = 1, \ldots, L$. Let $\Delta_2 > \Delta_1 > \max(x^*, \Delta_0)$. Since $\mathbf{R}^{\Delta_1} \subset \mathbf{R}^{\Delta_2}$, we have $I_{\Delta_1}^* \leq I_{\Delta_2}^*$. Note that the observation distributions involved in $S_{\Delta_2}^*$ are distributions on $[0, x^*]$ and hence $S_{\Delta_2}^*$ is in $\mathbf{R}^{\Delta_1}$ also. Hence $I_{\Delta_2}^* \leq I_{\Delta_1}^*$, which completes the proof of the theorem.

Remark

In view of Theorem 1.1, we see that the restriction that the observation distributions be distributions on $[0, \Delta]$ is not a serious restriction. However, if we do not assume $\Delta < \infty$, then for a policy S for which $Y_n(S) = \infty$ with probability 1 for some $n = n_0$, $I_N(S)$ will not be defined for $n > N_0$ and hence, our criterion of interest, $I(S)$, will not be defined.

ACKNOWLEDGEMENT

I thank Professor I. R. Savage for his constant encouragement during the course of this investigation.

REFERENCES

[1] Chitgopekar, S. S., "Continuous Time Markovian Sequential Control Processes," SIAM J. of Control 7, 367–389 (1969).

[2] Chung, Kai Lai, *Markov Chains with Stationary Transition Probabilities* (Springer, Berlin, 1966), 2nd edition.

CORRIGENDA

In the article "Alternate Methods of Project Scheduling With Limited Resources," (*NRLQ*, *20*, December, 1973, pp. 767–784) the horizontal lines depicting significant differences were unintentionally omitted from Tables 7, 10, 13, and 16. Corrected tables appear below. Also, the total sum of squares in Table 17 is 3,875,968, the total degrees of freedom in Table 5 should be 599, and the correct spelling of the third author in reference [2] is Shaffer.

TABLE 7. *Multiple Range Test for Total Project Delay*

Mean (In Days)	Treatment
1,502	GRRD RAN
1,487	GRRD
1,351	GTRD RAN
1,342	GTRD
1,206	GRRD RAN RES
1,201	GRRD RES
1,161	GTRD RAN RES
1,155	GRU RES
1,154	GRU RAN RES
1,151	LTF RAN RES
1,150	SIO RAN RES
1,138	SIO RES
1,132	GTRD RES
1,131	LTF RES
1,121	GRU
1,110	GRU RAN
1,054	LTF
1,043	LTF RAN
1,000	SIO RAN
998	SIO

TABLE 10. *Multiple Range Test for Weighted Total Delay*

Mean (In Man-Days)	Treatment
723,693	GRRD RAN RES
723,395	GTRD
723,127	GRRD RES
721,692	SIO
719,803	GRRD RAN
714,011	GTRD RAN
713,060	GRRD
710,927	SIO RAN
703,471	LTF RAN RES
703,235	GRU RAN RES
702,360	GTRD RAN RES
698,327	SIO RAN RES
697,270	GRU RES
695,129	SIO RES
693,777	GTRD RES
692,352	LTF RES
660,190	GRU RAN
655,869	LTF RAN
653,746	GRU
644,648	LTF

TABLE 13. *Multiple Range Test for Total Resource Idle Time*

Mean (In Man-Hours)	Treatment
68,203	SIO
67,694	SIO RAN
66,583	GRU RAN RES
66,571	SIO RAN RES
66,557	LTF RES
66,502	GTRD RES
66,484	GRU RES
66,419	GTRD RAN RES
66,414	SIO RES
66,394	LTF RAN RES
65,732	GRRD RES
65,681	LTF RAN
65,599	GRRD RAN RES
65,067	LTF
64,080	GTRD
63,987	GTRD RAN
63,595	GRU RAN
63,364	GRU
61,193	GRRD
60,736	GRRD RAN

TABLE 16. *Multiple Range Test for Computer Processing Time*

Mean (In Seconds)	Treatment
126	GRU RAN RES
120	GRU RES
107	GRU RAN
102	GRU
66	LTF RAN RES
65	LTF RES
48	LTF RAN
47	LTF
44	GRRD RAN RES
42	GRRD RES
42	GTRD RAN RES
40	GTRD RES
40	SIO RAN RES
38	SIO RES
23	GTRD RAN
22	GRRD RAN
22	GTRD
22	GRRD
21	SIO RAN
20	SIO

A horizontal line enclosing a group of means indicates that the means located within the group cannot be distinguished from one another at the 5% level of significance. Descriptions of each of these treatments are found in Table 3.

☆U.S. GOVERNMENT PRINTING OFFICE: 1974-541-390/4

INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.

NAVAL RESEARCH
LOGISTICS
QUARTERLY

JUNE 1974
VOL. 21, NO. 2
NAVSO P-1278

CONTENTS

OFFICE OF NAVAL RESEARCH
Arlington, Va. 22217